GENERATION OF AN RNA SENSOR TO PROBE GLYCOLYSIS IN LIVING CELLS VIA THE KEY METABOLITE FRUCTOSE 1,6-BISPHOSPHATE

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Ignazio Geraci

aus

Termini Imerese
Italien

Bonn 2021

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Günter Mayer

2. Gutachter: Prof. Dr. Michael Ryckelynck

Tag der Promotion: 06.05.2021

Erscheinungsjahr: 2021

A mia madre, mio padre, Lisa, e la mia amata famiglia.

Ed al principio inintelligibile che ci governa tutti.

ABSTRACT

Glycolysis is the cornerstone of cell metabolism and is one of the oldest biochemical pathways known to life. Yet, methods to study glycolysis in single living cells are lacking. The glycolytic intermediate Fructose 1,6-bisphosphate (FBP) generated interest because seminal studies correlated it with glucose influx. Here, a gene-encoded RNA fluorescence sensor was developed for the qualitative assessment of FBP to be used as a metabolic study tool.

For this purpose, the concept of allosteric communication of RNA modules was explored to convert the constitutively fluorogenic RNA Spinach into an FBP-dependent variant. Ad hoc designed libraries were equipped with Spinach's fluorogenic G-quadruplex core and tested to identify, within the RNA pools, whose sequence space would promote allosteric regulation. To scout FBP-responsive allosteric sequences, the RNA libraries were subjected to selection methodologies based upon SELEX (Systematic evolution of ligands by exponential enrichment) or on a microfluidic-assisted *in vitro* compartmentalization pipeline (µIVC). Numerous SELEX trials failed to enrich ligand-induced fluorescence sequences with either FBP or other test molecules, whereas µIVC allowed the identification of sequence 81 (s81), a ratiometric fluorescent sensor for FBP, from a chimeric library design of Spinach with the C45 aptamer, an *in vitro* selected RNA that binds to FBP.

S81 was characterized *in vitro* to assess its binding to FBP and DFHBI, the cognate dye of Spinach, and a possible sensing mechanism was laid out on the grounds of the parental chimeric library design. Finally, s81 was transiently expressed in embryonic kidney cells to probe the FBP variation upon oligomycin-induced glycolytic flux increase. Contrary to the literature, FBP was found to decrease upon oligomycin-treatment. Still, the assay provided a window to evaluate the sensor's performance in the cellular context by demonstrating s81 poly-specificity to phosphate-bearing metabolites.

Nevertheless, the identification of s81 as an FBP-responsive fluorogenic RNA showcased the advantages of combining structurally-guided designs with high-throughput screening pipelines to identify allostery-regulated iterations, which could be straight-forwardly characterized *in vitro* and tested in cells. This approach is preferable to current design-build-test cycles for RNA sensor discovery, for it allows the screening of entire libraries of millions of different sequences in a non-iterative and holistic manner.

ZUSAMMENFASSUNG

Die Glykolyse ist der Eckpfeiler des Zellstoffwechsels und zugleich einer der ältesten biochemischen Stoffwechselwege. Trotzdem besteht noch immer ein Bedarf an Methoden um die Glykolyse in lebenden Zellen zu untersuchen. Das Stoffwechselzwischenprodukt Fruktose 1,6-bisphosphat (FBP) ist in den letzten Jahren vermehrt in den Fokus der Wissenschaft gerückt, da nachgewiesen werden konnte, dass seine Konzentration mit der Aufnahme von Glukose korreliert. Es ist jedoch unklar, ob diese Korrelation in allen Lebewesen zu finden ist. In dieser Arbeit wurde der Ansatz aufgegriffen einen genetisch-enkodierbaren fluorimetrischen FBP Sensor für die qualitative Analyze von FBP Konzentrationen in lebenden Zellen als Werkzeug für die Stoffwechselforschung zu generieren.

Im Detail wurde dabei versucht mit Hilfe von allosterischen Wechselwirkungen von RNA Modulen die konstitutive fluorogene Eigenschaft des RNA Aptamers "Spinach" in ein FBP abhängiges RNA Sensormodul zu überführen. Dafür wurden Riboswitch-abgeleitete Bibliotheken verwendet, die mit dem fluorogenen G-Quadruplexkern des Spinach Aptamer ausgestattet waren, um diejenigen RNA Spezies zu identifizieren, die eine allosterische Regulation ermöglichen. Die Ermittlung der FBP abhängigen Sequenzen wurde mittels Methoden durchgeführt, die auf SELEX (Systematic evolution of ligands by exponential enrichment) sowie Mikrofluidikunterstützte in vitro Kompartimentalization (microfluidic-assisted in vitro compartmentalization, µIVC) basierten. Dabei stellte sich heraus, dass die mittels SELEX durchgeführten Versuche nicht erfolgreich waren, jedoch wurde mittels µIVC das Aptamer 81, ein ratiometirsch fluoreszierender Sensor für FBP, identifiziert. Aptamer 81 basiert auf einem chimären Bibliotheksdesign, das das Spinachaptamer mit dem FPB Aptamer C45 verbindet.

Die Bindungseigenschaften von Sequenz 81 (s81) wurden in Bezug auf seinen Liganden FBP sowie dem dazugehörigen Farbstoff DFHBI hin charakterisiert und ein möglicher Mechanismus auf dem zu Grunde liegenden parentalen Bibliotheksdesign wurde erarbeitet. Zuletzt wurde s81 transient in humanen embryonalen Nierenzellen exprimiert um die Änderung der FBP Konzentration zu messen ausgelöst durch oligomycininduzierte Steigerung des Glykolyseflusses. Dabei korrelierte FBP nicht mit der Oligomycinbehandlung. Jedoch bietet dieser Versuch einen Spielraum um die Sensorleistung im zellulären Kontest zu evaluieren durch die Demonstration der Polyspezifität von s81 gegenüber phophatbeinhaltenden Metaboliten.

Nichtsdestotrotz repräsentiert die Identifikation von s81 als ratiometrischer FBP-Sensor die Vorteile der Kombination aus strukturgeleitetem Design und Hochdurchsatzscreening zur Identifikation allosterisch regulierter Iterationen, die dabei direkt in vitro charakterisiert und in Zellen getestet werden können. Dieser Ansatz ist vorrangig geeignet für aktuelle Design-Aufbau-Test Zyklen für RNA-Sensor Entdeckungen, die ein Screening ganzer Bibliotheken von Millionen unterschiedlicher Sequenzen in einem nichtiterativen und ganzheitlichen Art und Weise erlaubt.

CONTENTS

List of Figures

LIST OF TABLES

LIST OF APPENDICES

List of Abbreviations and Acronyms

µIVC	Microfluidic-assisted *in vitro* compartmentalization
AC	Alternating current
AcCoA	Acetyl coenzyme A
ADP	Adenosine bisphosphate
AMP	Adenosine monophosphate
AQUA	Advanced quick-assembly
ATP	Adenosine triphosphate
cAMP	3′, 5′-cyclic monophosphate
Cra	*E. coli* transcription factor
ddH_2O	Double-distilled water
ddPCR	Droplet digital PCR
DFHBI	(5z)-5-(3,5-Difluoro-4-Hydroxybenzylidene)-2,3-Dimethyl-3,5-Dihydro-4h-Imidazol-4-One
DFHBI-1T	(5z)-5-(3,5-Difluoro-4-Hydroxybenzylidene)-2-Methyl-3-(2,2,2-Trifluoroethyl)-3,5-Dihydro-4h-Imidazol-4-One
DFHO	(5Z)-5-[(3,5-difluoro-4-hydroxyphenyl)methylidene]-3-methyl-2-(nitrosomethylidene)imidazolidin-4-one
DMEM	Dulbecco's modified Eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphates
dsDNA	Double-stranded DNA
DTT	1,4-dithiothreitol
ECAR	Extracellular acidification rate
eAMV RT	Enhanced Avian Myeloblastosis Virus reverse transcriptase
EtOH	Ethanol
$FADH_2$	Flavin adenine dinucleotide
FBP	Fructose 1,6-bisphosphate
FBS	Fetal bovine serum

FMP	Fructose 6 phosphate
FP	Fluorescent protein
FRET	Förster resonance energy transfer
FW	Forward-side primer
GFP	Green fluorescent protein
Glu	Glucose
Glu6P, G6P	Glucose 6 phosphate
GluN6P	Glucosamine 6 phosphate
GMP	Guanosine monophosphate
HBI	Hydroxybenzlidene imidazolinone
HCl	Hydrochloric acid
Hela	Henrietta Lacks cell-line
HEK293T	HEK293-derived cell line, expressing the SV40 large T antigen
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
iBMK	Immortalized Baby Mouse Kidney Epithelial Cells
KAc	Potassium acetate
LC-MS	Liquid Chromatography-Mass Spectrometry
NaCl	Sodium Chloride
NADH	Nicotinamide adenine dinucleotide
NaOAc	Sodium acetate
NGS	Next-generation sequencing
NHS	N-Hydroxysuccinimide
NMPs	Ribonucleotide monophosphates
NTPs	Ribonucleotide triphosphates
PCR	Polymerase chain reaction
PPase	Inorganic pyrophosphatase
PPP	Pentose phosphate pathway
PFK1	Phosphofructokinase enzyme
rcf	Relative centrifugal force

RNA	Ribonucleic acid
RV	Reverse-side primer
SAM	S-Adenosylmethionine
SELEX	Systematic evolution of ligands by exponential enrichment
SHAPE	Selective 2′-hydroxyl acylation analyzed by primer extension
SOC	Super Optimal broth with Catabolite repression
ssDNA	Single-stranded DNA
TCA	Tricarboxylic acid cycle
TPP	Thiamine pyrophosphate
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
tRNA	Transfer ribonucleic acid
WT	Wild-type

PREMISE

One century has passed since Otto Fritz Meyerhof's pioneering work on the Embden-Meyerhof theory of glycolysis, which promptly awarded him the Nobel Prize in physiology and medicine in 1923 at the astounding age of 39 years old.

His genius arises not merely per the identification of single steps of the glycolytic pathway but rather for describing – for the first time – that the steps of glucose degradation were coupled together to provide a function, that of storing, transforming, and efficiently releasing energy to the cell[1]. His work provided concrete evidence for the understanding of life processes as end-point stationary states, where energy and substances rest and flow at the same time.

To describe this phenomenon, Meyerhof liked to quote this poem from Conrad Ferdinand Meyer[2]:

Aufsteigt der Strahl und fallend gießt	*Up springs the spout and, falling, fills*
Er voll der Marmorschale Rund,	*To brim the marble basin's round,*
Die, sich verschleiernd, überfließt	*Which, under veiling, over spills*
In einer zweiten Schale Grund;	*Into a second basin's ground;*
Die zweite gibt, sie wird zu reich,	*The second one, too rich now, runs*
Der dritten wallend ihre Flut,	*Into the third its falling waves,*
Und jede nimmt und gibt zugleich	*And each one takes and gives at once*
Und strömt und ruht.	*And streams and stays.*

Together with Otto Heinrich Warburg, an old friend of Meyerhof and Nobel laureate famous for the "Warburg effect"[3], they laid the groundwork for modern biochemistry and oncology. At the very core of their work, we find glycolysis.

1 INTRODUCTION

1.1 The rediscovery of glycolysis with a single-cell focus

1.1.1 The relevance of glycolysis in cellular physiology and cancer

Glycolysis is the bedrock of almost all living systems' metabolism[4]. By providing energetic and biosynthetic intermediates, glycolysis lays the groundwork for other secondary metabolic pathways to occour[5]. The glycolytic pathway comprises ten enzyme-catalyzed reactions that convert one molecule of glucose into two molecules of pyruvate (**Figure 1**)[6,7].

From the energetic standpoint, glycolysis supplies NADH, ATP – central energy storage molecules[8] – and pyruvate that enters the tricarboxylic acid (TCA) cycle within mitochondria to produce more NADH and also $FADH_2$[9]. These reducing agents fuel the mitochondrial electron transport chain of oxidative phosphorylation, thereby coupling glycolysis with mitochondrial respiration. Respiration is the primary source of ATP in eukaryotic cells but, conversely to glycolysis requires oxygen[9]. Therefore, cells dynamically switch between glycolysis and respiration according to environmental circumstances, such as normoxia and hypoxia, or energetic demand[10].

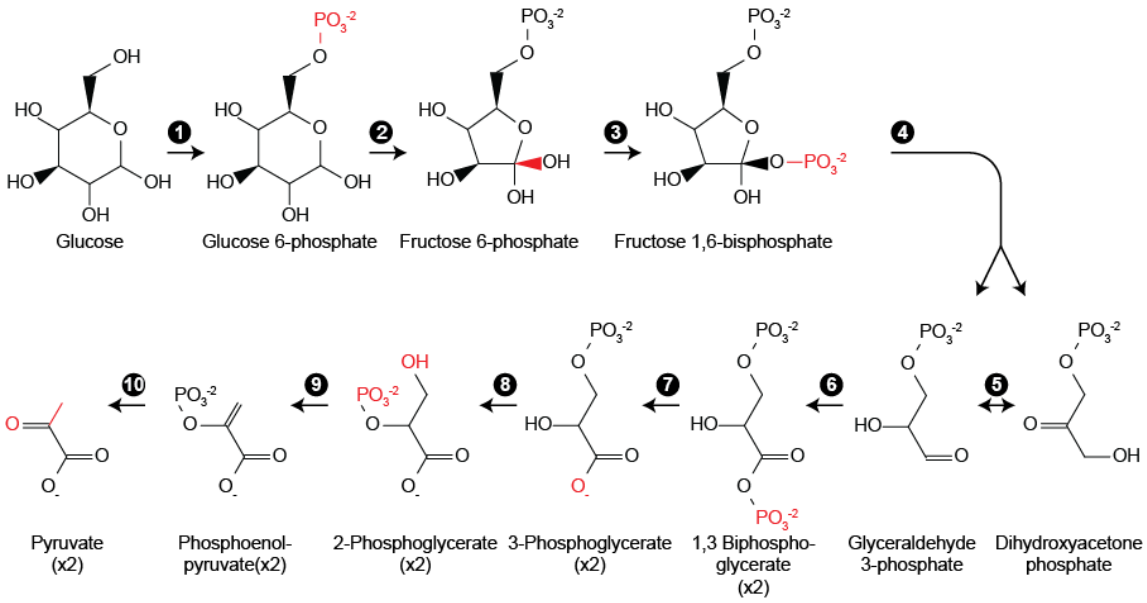


Figure 1 | The pathway of glycolysis.
The chemical progression from glucose to pyruvate within the glycolysis pathway: the chemical modifications are highlighted in red. To note, while phosphorylation 1 and 3 are ATP-consuming, de-phosphorylation 7 and 10 release ATP molecules, with a net production of 2 ATPs per glucose molecule.

From the biosynthetic standpoint, glycolysis produces the substrates required for cell growth and proliferation. For instance, glucose 6-phosphate branches out of glycolysis and enters the pentose phosphate pathway (PPP), which supplies metabolites for nucleotide biosynthesis, and more broadly, for DNA replication and RNA transcription[9]. Another example is 3-phosphoglycerate, which supplies one-carbon metabolism that is a pathway connected to several biosynthesis outputs[11].

In mammals, glycolysis regulation was investigated in the physiological processes of cell proliferation[12], immune activation[13], and angiogenesis[14]. More importantly, it becomes vital to investigate the causes of a lack of glycolysis regulation, such as in cancer[15].

The alteration of glycolysis typically seen in cancer cells is referred to as the Warburg effect[3], which causes cells to boost their glycolytic rate independently from

oxygen[3] or mitochondria functionality[16]. This so-defined “reprogramming” of glycolysis supports the acquisition and maintenance of malignant properties, making it a hallmark of cancer[17]. For instance, cancer cells can proliferate at a pathologically elevated rate because of the enhanced biosynthetic rate and can withstand survival in stressful environments due to the altered bioenergetics, which gives them an advantage of survival over normal cells[17,18].

Despite its relevance for cancer physiology, the function of the Warburg effect, and, more importantly, the clear understanding of how to exploit it for therapeutic benefit is still a matter of debate[19]. A confounding reason is that many cancer cell lines still possess functional mitochondria, and can shift to oxidative phosphorylation to meet their ATP demand[20,21], even reversibly from one pathway to the other[22]. Interestingly, a recent model proposed that “glycolytic” cells within a tumor could establish a metabolic symbiosis with the “oxidative” ones based on the shuttling of lactate[23], closely resembling metabolic heterogeneity studies performed on microbial populations[24].

Naturally, most cancers are the product of overlapping factors[25], and approaching them solely from the glycolytic paradigm might be too simplistic[26]. Nevertheless, the development of novel methods for the investigation of glycolysis in single living-cells is crucial to comprehend how cell-to-cell metabolic variability and cancer insurgence intertwine and, encouragingly, would pave a new way to study also physiological metabolism[27].

1.1.2 Cell-to-cell heterogeneity on the metabolic level

The finding that metabolic heterogeneity plays a vital role in prokaryotic and eukaryotic clonal populations alike has set the stage for the rediscovery of metabolic research with a new single-cell focus[28]. The experimental evidence that fuelled the interest towards metabolic heterogeneity was the notion that there is an element of stochasticity in every cellular process that causes cells, even with the

same genetic and epigenetic expression pattern, to display different phenotypes[24,29].

Metabolic heterogeneity can arise in response to changing environments, where cells start cooperating to minimize the stress of the colony. Interestingly, yeast colonies[30] and cancer cells[23] have been found to cooperate by producing and exchanging different metabolic precursors, thus creating metabolically-specialized subpopulations and increasing the fitness of the population.

However, metabolic heterogeneity is not always beneficial, as it was observed in yeast cells in response to dynamic glucose change[28]. The variation of feeding glucose in the growing medium causes a metabolic subpopulation of non-growing cells to arise due to malfunctioning glycolytic regulation. The establishment of a non-growing cellular subpopulation due to altered metabolism has particular relevance to biotechnological applications, e. g. the decrease in production yield obtained while scaling-up a fermentation process[28].

Sources of metabolic heterogeneity can also be the cell cycle or aging, which caused unsynchronized temporal behavior of the individual cells and was observed in cycle-synchronized yeast cells[31] and aging budding yeasts[32] upon every cell division.

1.1.3 From traditional to gene-encodable approaches to probe glycolysis

Regarding glycolysis research, the current approaches estimate glycolytic flux – namely the uptake of carbon molecules through the glycolysis pathway – by measuring the glucose-to-lactate conversion or by the extracellular acidification rate (ECAR) of cultured cells[9].

Glucose and lactate are the two extracellular end-points of glycolysis and can thus be easily measured by commercially available kits[9] coupled with standard lab spectrophotometers. Instead, measuring ECAR requires specific equipment[9], and

determines glycolysis from the acidification of the culture medium resulting from the excretion of lactic acid per unit of time. Mainly, ECAR has been used in combination with oxygen consumption to measure glycolysis and oxidative phosphorylation simultaneously[33].

Glucose-to-lactate and ECAR provide a reasonable estimation of the glycolytic flux, are relatively easy to perform and cell non-destructive[9], but are population-averaging techniques and cannot be applied to unravel cell-to-cell heterogeneity on the metabolic level. One methodology which holds great promise for uncovering metabolic heterogeneity is single-cell metabolomics by mass spectrometry, but it has the downside of being cell-disruptive, and it is not suitable for dynamic measurements[34,35].

Alternatively, the construction of a genetically-encoded biosensor could allow the interrogation of glycolysis dynamically, intracellularly, and in living single-cells, thereby meeting all the requirements for metabolic heterogeneity studies. Gene-encoded biosensors link the fluctuation of a specific cellular analyte to a reporter domain, mostly based on fluorescence, and enable monitoring at the cellular level with precise temporal resolution[36]. These sensors are based on fluorescent proteins[37] or, although more recently, on fluorogenic RNAs[38]: pros and cons of both design strategies will be discussed in the subsequent sections (Section 1.1.5 and 1.2). For metabolic research, there have already been several applications for gene-encoded biosensors from enzyme engineering[39,40] to cell factory optimization[41,42] and real-time metabolite monitoring[38]. The biotechnological application of gene-encoded biosensors showcases their value as analytic tools to provide insights into metabolic networks and pathway regulation, making it possible to study biochemistry *in vivo*[36].

1.1.4 FBP as the target for biosensor construction

Fructose 1,6-bisphosphate (FBP) is an excellent target for constructing a biosensor with relevance to glycolytic flux regulation. FBP is produced by the third enzymatic

step of glycolysis (**Figure 1**) by the enzyme phosphofructokinase (PFK1), tightly regulated in glycolysis[43]. PFK1 is allosterically modulated by the ATP/AMP ratio to provide a negative feedback mechanism when the cellular ATP level is high[9]. Due to its tight regulation, FBP is the first intermediate metabolite of glycolysis at which glucose carbons become fully committed to glycolysis[9]. Seminal studies were conducted to correlate intracellular FBP level to the glycolytic flux in *S. cerevisiae*[44-46], *E. coli*[47,48], and cultured mammal cells[15].

In *S. cerevisiae*, there is a substantial body of evidence showing how FBP mirrors glycolytic flux both in glucose batch and glucose-limited chemostat cultures[44]. Furthermore, when glucose is added as a carbon source in a minimal culture, a corresponding increase in glucose influx and the FBP concentration is observed, indicating that FBP reflects the glycolytic flux both dynamically and in steady-state[45]. The correlation between FBP levels and glycolytic flux in *S. cerevisiae* was recently enabled by the increasing availability of strains engineered for specific metabolic pathways, e. g., the TM6* strain[46]. TM6* carries one chimeric hexose transporter instead of 17 like the wild-type (WT), causing a low glucose uptake, low glycolytic flux, and a low intracellular concentration of FBP. Interestingly, the glycolytic flux can be re-established to the WT condition by growing the strain on maltose, carried into the cell by a different transporter[44,46]. The TM6* strain, in particular, was used to establish an *in vivo* system able to generate a range of glycolytic fluxes, which were used to evaluate the flux-reporting capacity of a sensor protein for FBP[46].

In *E. coli*, Kochanowski *et al.* suggested that the glycolytic flux information imprints into the intracellular level of FBP and is consequently read out by protein effectors, e. g., the transcription factor Cra that is negatively regulated by intracellular FBP[47,48]. FBP-Cra flux sensor components were correlated to the glycolytic flux using ^{13}C flux analysis on glucose-limited cultures of *E. coli,* demonstrating that the FBP-Cra axis functionally responds to flux variation. Correctly, intracellular FBP increases with glycolytic flux, while the activity of Cra decreases accordingly[47].

Nevertheless, the authors did not investigate direct causation between the increase of FBP and the downregulation of Cra transcriptional activity.

For mammalian cells, the most extensive evidence of the relationship between intracellular FBP and the glycolytic rate was obtained by Tanner *et al.* on mouse kidney and fibroblast cells[15]. The limiting steps of the glycolytic flux were identified by systematically over-expressing each enzyme of glycolysis and by measuring the corresponding consumption of glucose and secretion of lactate, giving an estimate of the glycolytic influx and efflux (Section 1.1.3). Among all the overexpressed genes, PFK1 was reported together with genes associated with glucose import and lactate export, concluding that the glycolytic rate is mirrored intracellularly by the concentration of FBP.

The FBP level in cells seems to step above as a marker of glycolytic flux in prokaryotes[47,48], single-cell eukaryotes[44], and even complex eukaryotes[15]. However, the cellular detection of FBP requires new molecular tools to confirm whether this correlation holds across all organisms. The development of a gene-encodable FBP sensor would be well-suited for testing the FBP-glycolytic flux correlation in living cells.

For those organisms where the FBP-glycolytic flux correlation is confirmed, the measurement of intracellular FBP would allow estimating the glycolytic flux variability within a cell population and, consequently, generate exciting possibilities for answering pressing metabolic research questions that require probing glycolysis in single-cells.

1.1.5 Corrival approaches for cellular detection of FBP

The design of fusion-based design comprising fluorescent proteins (FP) such as the green fluorescent protein and derivatives is a core methodology for developing encodable fluorescent sensors[37,49].

FPs-based sensors can report and produce a ratiometric signal in response to a specific ligand by expressing them under the control of a ligand-binding

transcription factor[46,47] or in a cis-regulatory RNA element named riboswitch[50]. A recent study has used the transcription factor CggR of *B. subtilis* as a sensor for FBP, demonstrating the applicability of this system in flow cytometry and time-lapse fluorescence microscopy in yeast cells[46]. Transcription-based protein sensors allow non-disruptive measurements of metabolites in living cells. However, these sensors lack temporal resolution due to the lag-time between the ligand variation and the required time to translate and mature the nascent fluorescent protein[51]. Similar to single-cell mass spectrometry, transcription-based protein sensors cannot report real-time changes in the analyte concentration.

A significant step towards a real-time metabolite sensor was obtained by Förster resonance energy transfer (FRET)-based FP sensors. FRET relies on the energy transfer between two interacting FPs with overlapping emission and excitation spectra that come in proximity to one another and is the methodological base of FRET fluorescence microscopy[52]. Since the FRET signal strongly depends on distance, chimeric designs of FPs, and ligand-binding proteins that point at obtaining a structural variation of the relative distances between the energy-donor and energy-acceptor FPs in response to binding the target ligand enable to engineer protein-based sensors that dynamically respond to a given ligand. FRET sensors are available for essential components of central carbon metabolism, such as ATP[31] and pyruvate[53], but their reduced signal-to-noise ratio, combined with the limited pool metabolite-specific protein domains[54], limits their applicability.

In addition to their limitations, the application of protein-based sensors is disadvantageous for imaging metabolite dynamics and, more broadly, for metabolomics analysis. The expression of a protein-based sensor heavily exploits the biosynthetic translation machinery, thereby depleting the cytoplasm of the necessary metabolites and co-factors[55], primarily when they rely on viral promoters to generate high amounts of protein. While this issue might not be problematic for biotechnological application, for glycolysis-focused studies, the prospect of

measuring metabolism with a sensor that causes a heavy metabolic burden to the host cells is an untenable one[56,57].

1.2 RNA-based fluorescence sensors

Compared to the protein-based counterpart, RNA-based fluorescence sensors present enormous potential as metabolite-tracking tools. Firstly, their expression influences the hosts' metabolism to a lesser extent because RNA does not rely on the cells' translational machinery, overcoming the problems inherent to protein-based expression (Section 1.1.5). Secondly, target-interacting RNA aptamers can be identified by *in vitro* selection techniques (Section 1.2.1), together with fluorogenic RNA scaffolds (Section 1.2.2). Finally, their advantageous *in vitro* assembly properties allow the design and rationalization of structure-switching RNA sensors (1.2.3).

1.2.1 RNA aptamers and SELEX

RNA aptamers are oligonucleotide sequences that fold into a specific 3D architecture and can bind a specific ligand[58,59]. Aptamers can also be structured single-strand DNA sequences, but, albeit having other appealing advantages[60], they are not readily expressible in a host cell and thus inappropriate for metabolite monitoring. The applications of RNA aptamers for cellular sensing are far-reaching and tethered to the cellular component they target, e. g. proteins[61,62], small molecules[63,64], ions[65,66] or even whole cells[67]. Sequences that exhibit affinity towards a specific target can be identified inside the scaffold of naturally-occurring riboswitches[50] or selected *in vitro* amongst a pool of random sequences in a cyclic process of binding-retention-elution named Systematic Evolution of Ligands by Exponential Enrichment (SELEX)[68,69].

SELEX retains and propagates ligand-binding RNA sequences of an RNA pool in a process akin to Darwinian evolution[68], but using the advantage of enzymatic reactions that can be performed in an Eppendorf tube. Not requiring advanced

equipment, SELEX is a straight-forward methodology applied by research labs throughout the world during the last 30 years[70], creating an extensive repository of ligand-specific aptamers[71]. Since the conformational structure of aptamers strongly depends on ions, temperature, and pH, the selection of aptamers for cellular application is commonly performed in buffer systems resembling the inner composition of cells[72].

1.2.2 The revolution of light-up RNA aptamers

In the late '90s, Wilson and co-workers performed SELEX on the two small organic molecules malachite green[73] and sulforhodamine B[74]. The resulting RNA aptamers exhibited a novel, emergent property of enhancing the fluorescence of the normally nonfluorescent triphenylmethane dyes[75]. However, the two aptamer-dye pairs could not be applied as they exhibited dye toxicity and high background fluorescence. From this example, a rush occurred to establish the first fluorogenic RNA system applicable in living cells[76].

The first RNA-dye system to set the benchmark for all later innovations in the field was Spinach[77]. Spinach is an umbrella term for a series of RNA aptamers that bind a chemically synthesized variant of the 4-hydroxybenzlidene imidazolinone (HBI), the active fluorophore of GFP[77]. Many variants of this dye interact with Spinach *in vitro*, creating an attractive palette of RNA-fluorophore complexes. However, DFHBI (2,5-difluoro-4-hydroxybenzlidene imidazolinone, chemical structure in **Figure 8 a**) was the only one applicable in bacterial and mammalian cell cultures for its cell-permeability, nil cytotoxicity, and bright green fluorescence in complex with Spinach RNA[77].

When first published, Spinach was readily applied for live-cell imaging of 5S ribosomal RNA via a simple RNA-tagging strategy that displayed the potential of fluorogenic RNAs[77]. Subsequently, imaging of RNAs through Spinach-tag was employed in bacteria[78,79], yeast[80], viruses[81,82], mammalian cells[83], and even algae[84] to demonstrate the bio-orthogonality and broad applicability of the system. Spinach

was also employed to track RNAs *in vitro*[85] and allowed to couple transcription and translation rates of an encoded protein in cell-free protein expression[86,87]. The potential of Spinach as an RNA reporter was shown in the engineering of next-generation genetic circuits and logic gates[88,89]. Another noteworthy application of Spinach was in drug screening, e. g. of inhibitors for RNA polymerase[85] and RNA-modifying enzymes[90]. In particular, Spinach was revolutionary to scale-up the throughput of a screening system from microtiter plates[85] to high-throughput[90] and even ultrahigh-throughput workflows[91].

Nonetheless, the first version of Spinach was far from an optimal imaging tool[83]. The RNA-dye complex suffered from limited brightness and poor photostability in live-cells, which could only be partially overcome by using tandem repeats of Spinach-tag[79] or by optimizing the imaging acquisition settings[92]. Novel improved RNA variants of Spinach were identified by systematic mutagenesis[83,93], miniaturization[94], and improved folding in cellular milieu[95].

Parallelly, other fluorogenic RNA-dye pairs were identified by SELEX that partially overcome the limitations of Spinach-DFHBI. Noteworthy examples are Corn[96,97], o-Coral[98], Chili[99], and the series of RNAs known as Mango[100-102] and SiRA[103]. Although the novel fluorogenic RNAs demonstrated significant improvements in brightness, photostability, and tuned-in spectral properties, Spinach remains the most well-established and characterized system nowadays.

1.2.2.1 The structure and photophysics of Spinach

Since the first publication of the Spinach fluorogenic RNA system, plenty of insights into the mechanism and photophysics of the fluorescence activation of DFHBI were obtained. First and foremost, the resolved crystal structure was published by two independent research groups[94,104] using Spinach 24.2, the original aptamer identified through SELEX by Paige *at al.*[77].

In both resolved structures, Spinach folds into a coaxial helical stack (**Figure 2**), contradicting the computational predictions initially used to establish the

presumptive aptamer conformation[77]. The stack is discontinued by two irregular junctions, one is expendable (J1-2, **Figure 2**), while the other– named J2-3 – constitutes the core domain that coordinates with DFHBI. The core of J2-3 comprises two G-quartets staked above a supporting mixed tetrad and structurally stabilized by potassium ions, constituting the Spinach G-quadruplex domain (**Figure 2**). The top G-quartet of Spinach interacts with the DFHBI fluorophore, which also displays a planar conformation when bound (**Figure 3**). DFHBI binds in the planar conformation on top of the G-quadruplex domain and is stabilized by π-π stacking interactions with the top tetrad (**Figure 3**).

The stabilization of the dye's planar conformation creates a highly fluorogenic complex. Upon excitation with blue light, DFHBI cannot dissipate its energy through non-radiative decay when bound in this conformation and uses the radiative decay of fluorescence instead, thereby increasing the quantum yield of the complex significantly[105]. DFHBI displays fluorescence excitation maxima of 447 nM alone, 460 nM when bound to Spinach, and a peak fluorescence emission of 501 nM[77]. The fluorophore is further stabilized by an unpaired Guanine residue and a planar U-A-U bases triple, creating a binding pocket on top of Spinach G-quadruplex (**Figure 3**).

The binding pocket of Spinach creates a quasi-hydrophobic environment that shields DFHBI from interacting with water molecules, in that resembling the structure of GFP fluorophore[106]. For instance, the fluorophore HBI is deeply buried into GFP and is stabilized by van der Waals interaction with aliphatic residues that restrain and induce the fluorophore's fluorescence[106,107]. In contrast, Spinach's π-staking interactions only partially shield DFHBI from interacting with the solvent, creating an open structure where the fluorophore associates and dissociates[94,104].

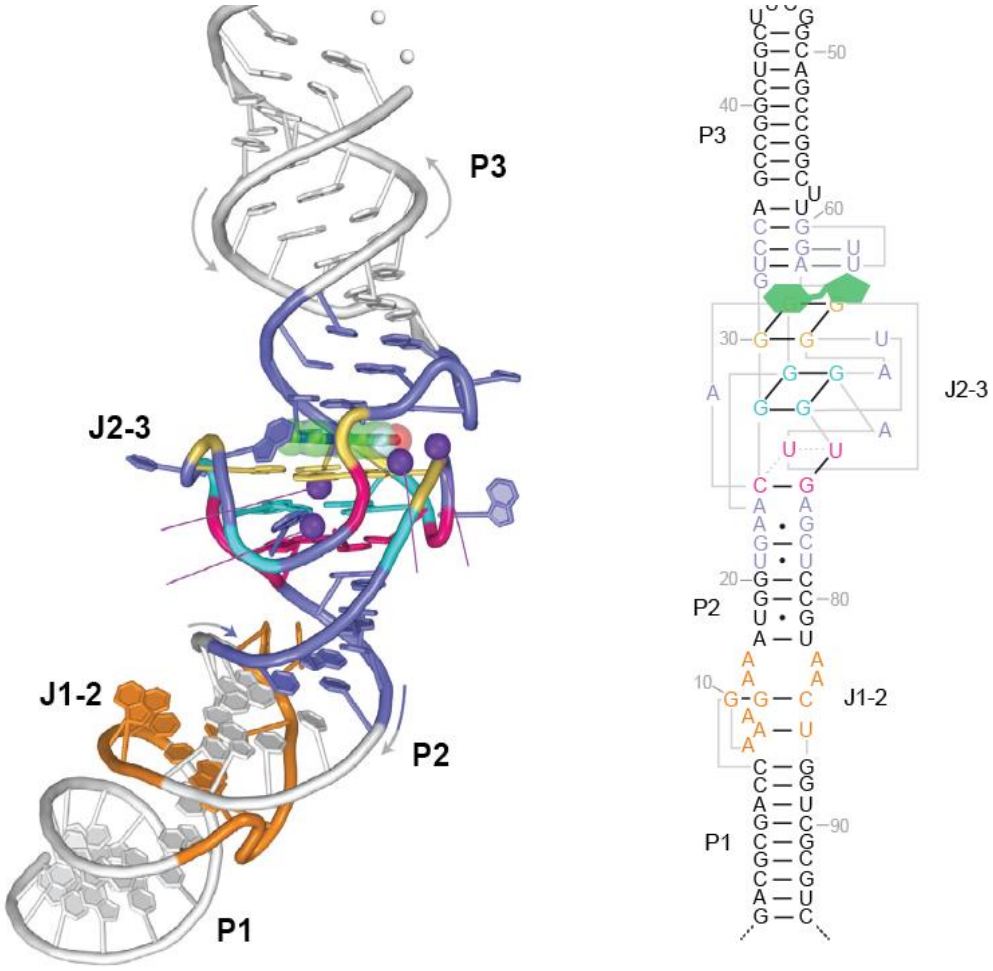


Figure 2 | Structure of Spinach aptamer in complex with DFHBI.
a, Cartoon representation of the resolved crystallographic structure of the first version of the aptamer, Spinach 1.2[77], published in Warner *et al.*[94] The P1, P2, P3 duplexes (white), the L3 loop (white), the two irregular junctions J1-2 (orange), and J2-3 (blue) domains of Spinach are highlighted. In particular, the three tetrads of J2-3 are color-coded to differentiate the two G-tetrads (yellow, cyan) and the mixed supporting tetrad (pink). Purple spheres represent K^+ ions. **b,** Sequence of Spinach 24.2 folded and color-coded according to **a.**

The loose interactivity between Spinach and DFHBI is displayed in the relatively fast fluorescence decay of the complex and the susceptibility to photobleaching[108]. Experimentally, the fluorescence lifetime of Spinach-DFHBI is ~4 nanoseconds, 1.5 times longer than GFP (2.7 ns)[92]. However, the Spinach-DFHBI complex also

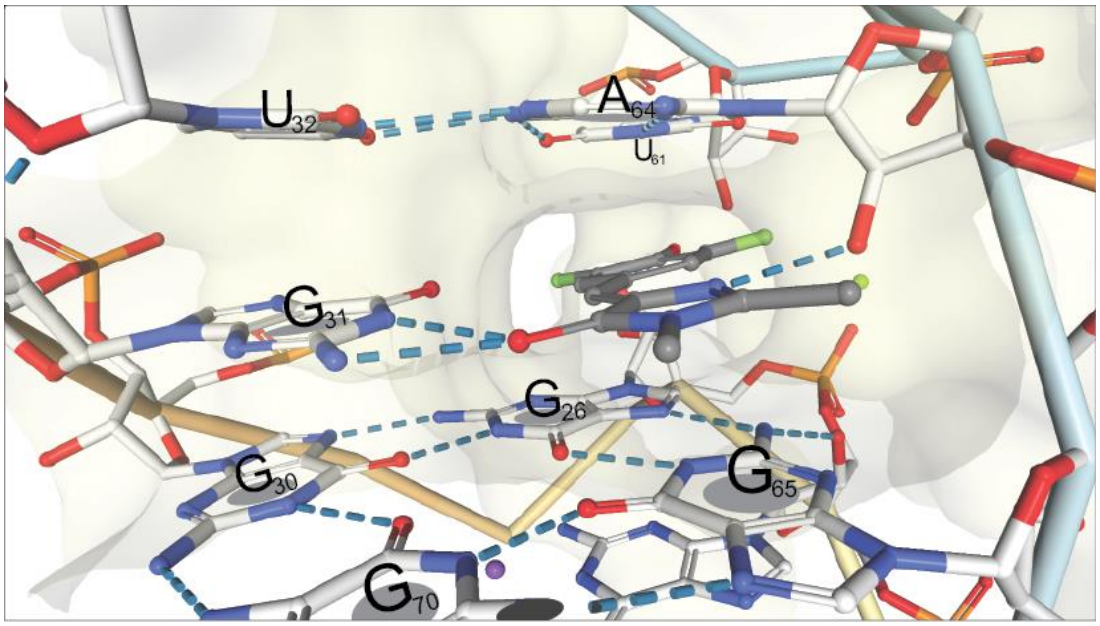


Figure 3 | Crystal structure of the binding pocket of Spinach bound to DFHBI.
Carbon atoms of DFHBI are drawn in grey, those of nucleotides in white. Nitrogen atoms are drawn in blue, oxygen in red, and fluorine in green. DFHBI is sandwiched between the top G-quartet tetrad formed by nucleotides G26, G30, G65, G70, and the planar U32-A64-U61 bases triple. The hydrogen-bonds between the residues are shown. Hydrogen-bonds stabilize the fluorophore to the DFHBI carbonyl oxygen and the unpaired G31, also depicted. The structure was handled with the PDB web software (PDB code 4T62).

displays a remarkable fluorescence signal recovery, whereas GFP undergoes irreversible photobleaching that causes the signal to drop[92]. Nevertheless, the recovery of the complex is hindered by DFHBI photoisomerization from the *cis* to *trans* conformation[108]. The *trans*-DFHBI is less stable and reported to have a dimmer fluorescence compared to the *cis*- photoisomer in complex with Spinach.

The photostability of Spinach was improved in combination with DFHBI-1T, a fluorophore variant with optimized spectral properties and enhanced brightness compared to DFHBI[109]. DFHBI-1T allowed RNA imaging in mammalians without the need for background fluorescence subtraction, which was necessary with Spinach-DFHBI[109].

Likewise, the RNA structure of Spinach was optimized from the original 97 nt sequence 24.2 identified by SELEX[77] (**Figure 2**). Jaffrey and co-workers initially identified Spinach 2 by systematic mutagenesis[83]. Spinach 2 exhibited an increase in folding and thermal stability in living cells, but the performance of the two RNAs was comparable[83,110].

Successively, they identified the Broccoli aptamer using a combination of SELEX and fluorescence-based selection[111]. Broccoli is a 49 nt RNA aptamer that also binds DFHBI-related dyes like Spinach but displays a robust fluorescence and improved folding in low-cytosolic magnesium concentration, which even Spinach 2 suffers from[111-113]. However, seminal comparative analysis rebutted the claim that Broccoli is a better alternative than Spinach for cellular application[114,115]. Parallelly, Ryckelynck and co-workers published an improved version of the Spinach aptamer, named iSpinach[114]. Like Spinach 2, iSpinach aptamer was identified by mutagenesis but through a high-throughput screening methodology, which uses microfluidic-assisted *in vitro* compartmentalization[116]. The iSpinach aptamer exhibits appealing properties such as increased brightness, higher thermal stability, and optimized salt-sensitivity, which, however, were only tested *in vitro*.

Last but not least, the crystal structure obtained from Spinach 24.2 was used to construct Baby Spinach, a miniaturized version of the aptamer lacking the expendable J1-2 junction and shortened coaxial helical stacks[94]. The miniaturized Baby Spinach aptamer displayed the best cellular performance in folding and brightness compared to all previous Spinach versions[115], especially in combination with structural scaffolds[112]. Baby Spinach became the benchmark aptamer for the Spinach RNA series, and it is still applied nowadays[117,118], albeit improved fluorogenic RNA systems, which are now available[96,98,99,103].

1.2.3 Design of structure-switching RNA

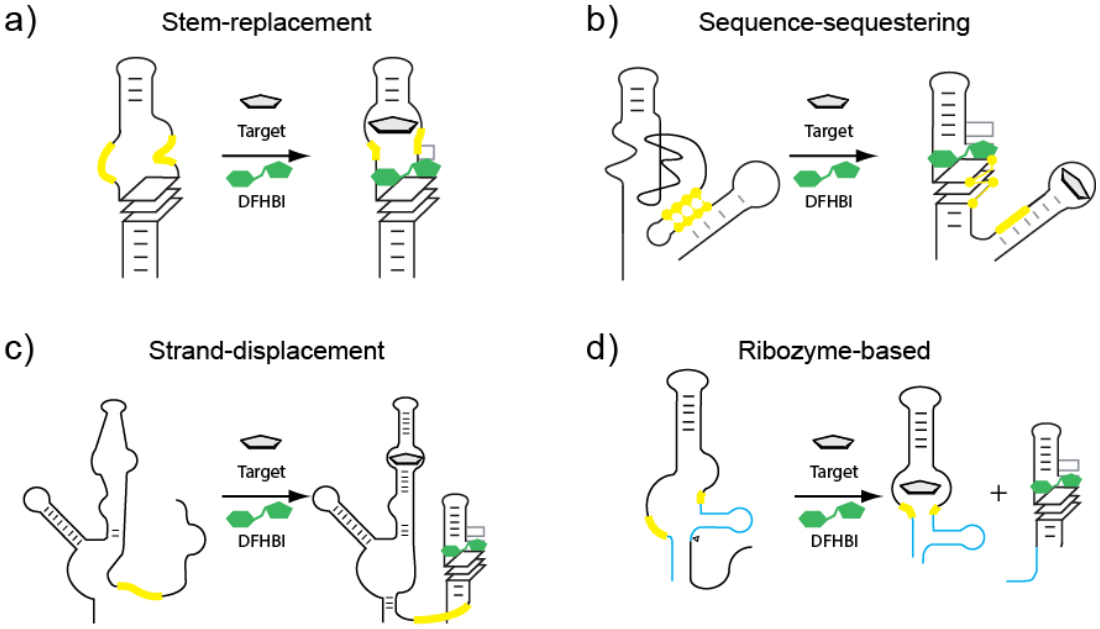


Figure 4 | Strategies for the rational design of allostery-controlled RNA elements. The main strategies for the design of RNA sensors are schematically presented. **a,** The stem-replacement allows to obtain allosteric control of the sensing domain over the reporter domain by substitution of the critical stem of the reporter domain with the target-specific aptamer. **b,** Sequence-sequestering approach is an alternative method to **a**, and conjoins to aptamers in-frame to sequester a critical stretch of the reporter aptamer, which will be released upon target-binding. **c-d,** Other effective designs, but lesser applied, unfold the reporter domain by forcing it into an open conformation, which is then released by strand-displacement (**c**) or by the activity of an embedded ribozyme (highlighted in blue, **d**). A meeting point between all different strategies is using a communication module (highlighted in yellow), which prime the RNA modules to collaborate allosterically.

The availability of a defined crystallographic structure renders Spinach a particularly well-suited RNA aptamer for engineering a fluorescent, allosteric sensor. The clear-cut relationship between each nucleotide of the structure and the end-point fluorescence of the system were critical to approaching an emergent property like allostery in a rational manner. Thence, Spinach-DFHBI-based RNA sensors were soon published with the capability of detecting microRNAs[119,120] or proteins[91,121] both *in vitro* and in cells.

Metabolite	Fluorogenic RNA	Design strategy	Application	Year[REF]
Adenosine	Spinach 1.2	Stem-replacement	*in vitro*	2012[122]
ADP	Spinach 1.2	Stem- replacement	*E. coli*	2012[122]
SAM	Spinach 1.2	Stem- replacement	*E. coli*	2012[122]
GTP	Spinach 1.2	Stem- replacement	*E. coli*	2012[122]
c-di-GMP	Spinach 1.2	Stem- replacement	*E. coli*	2013[123]
c-di-AMP	Spinach 2	Stem- replacement	*Listeria, E. coli*	2015[124]
TPP	Spinach 2	Sequence- sequestering	*E. coli*	2015[125]
Guanine	Spinach 2	Sequence- sequestering	*in vitro*	2015[125]
Adenine	Spinach 2	Sequence- sequestering	*in vitro*	2015[125]
SAM	Spinach 2	Sequence- sequestering	*in vitro*	2015[125]
2',3'-cGAMP	Spinach 2	Stem- replacement	Mammalian extract	2016[126]
Glycine	Spinach 2	Strand- displacement	*in vitro*	2016[127]
c-di-GMP	Spinach 2	Stem-replacement	*E. coli*	2016[128]
5HTP	Broccoli	Stem-replacement	*E. coli*	2017[129]
Theophylline	iSpinach	Stem-replacement	*in vitro*	2019[130]
Theophylline	Broccoli	Ribozyme-based	*E. coli*	2019[131]
SAM	Corn	Stem-replacement	HEK293T	2019[132]

Table 1 | List of designed RNA-based fluorescent sensors developed for metabolite ligands.

Spinach-based fluorescent sensors were mainly applied to visualize the presence and absence or dynamic changes of the concentration of small-molecules and metabolites, for which molecular tools are highly demanded, but remain elusive to identify with other classes of biomolecules (Section 1.1.5)[54]. A detailed list of all available light-up sensors for small molecules is depicted in **Table 1**. To obtain a

responsive sensor, the fluorogenic RNA is conjoined to a ligand-interacting RNA aptamer, previously identified by SELEX[122,133] or, instead, isolated from the scaffold of a riboswitch[125,127]. Strategies that were applied to fuse these two RNA modules are summarized in **Figure 4**.

Among all different design strategies, the motif that unites the two sequences into a unique functional scaffold plays a crucial role. This motif has been called in different ways: transducer[122], communication module[133], or actuator[127]. Nevertheless, this motif has the complementary role of transducing the structural rearrangement of the aptamer module to the fluorogenic module, thus attaining allosteric control of the former towards the latter.

Most biosensors were constructed by *stem-replacement* (**Figure 4 a**), thereby substituting the coaxial helix P3 of Spinach (RNA structure in **Figure 3**) with the minimal sequence motif of the ligand-interacting aptamer. In this way, the J2-3 domain correctly folds and interacts with DFHBI only when the ligand is embedded within the aptamer region, therefore emitting fluorescence ratiometrically. Noteworthy examples are the sensor for the cyclic dinucleotide c-di-GMP and c-di-AMP, a class of signaling molecules that controls a wide variety of pathogenic responses in bacteria[123,124,128].

When the aptamer domain is not readily applicable for this design, e. g, because the ligand-binding pocket resides in a loop region, *sequence-sequestering* is another viable strategy. In these designs, the two modules are conjoined in-frame to sequester a critical stretch of the reporter domain (**Figure 4 b**). The addition of the ligands allows the reporter domain to fold and switch to its fluorescent conformation. This strategy was successfully applied to convert a natural TPP-sensing riboswitch into a TPP fluorogenic biosensor[125].

Strand displacement is an alternative design strategy that exploits the structure-switching mechanism of riboswitches (**Figure 4 c**) and embeds the structure of a fluorogenic RNA within the riboswitch scaffold. In this design, the fluorogenic RNA

is forced into an open conformation by a displacement strand, which is then released by the addition of the ligand. Albeit being applied for small molecules, this strategy was more useful to devise microRNA sensors[119,120]. Finally, the *ribozyme-based* approach (**Figure 4 d**) embeds the catalytic core of a ribozyme to a ligand-sensing aptamer to release the light-up RNA upon the binding of the given ligand. As the cleavage is not reversible, this particular strategy is not suited to visualize metabolite dynamics but could be employed to detect low-abundance metabolites using the signal accumulation of ribozyme-cleaved light-up RNAs[131].

Although still in the pioneering age, given that the first RNA-based fluorescence sensor date to 2012[122], these sensors were straight-forwardly applied *in vitro*, e. g. to develop a fluorescence-based assay to screen enzyme activities in high-throughput[134]. Several proofs-of-concept were also applied in cells, mainly in bacteria, to showcase the capability of these sensors to report dynamic changes of a given molecule in living cells (**Table 1**).

Initially, the application of these RNA-based fluorescent sensors in mammalian cells was limited due to the poor folding and brightness of these RNA structures in the cellular context. However, expressing the RNAs as circular constructs seems to significantly extend their performance and relative concentration in the cytosol, resulting in a robust fluorescence comparable to the fluorescence protein-counterpart[135]. The integration of this circular expression methodology with the newly discovered RNA-dye pair Corn-DFHO enabled the design of a SAM-biosensor variant that could be finally applied to visualize SAM dynamics in living mammalian cells[**132**], thus opening a new chapter in the field of bioimaging.

1.3 The other side of the coin: structure-switching aptamers

Identifying a suitable RNA aptamer specific for the desired ligand is another matter that must be carefully taken into account when developing an RNA-based fluorogenic sensor. While the repertoire of small-molecule interacting RNA aptamers is vast[71], the uncomfortable truth is that most of these sequences cannot

be utilized for this purpose. Hammond quite simply crystallized the reason in her recent review on RNA-based fluorescent sensors: “Most *in vitro* selected aptamers may not be suitable as binding domains for the same reason that antibodies are not used in protein-based biosensors: they are selected for binding, not conformational change”[38].

As all directed evolution approaches, SELEX provides its results by following a path of least resistance, thus only retaining those sequences that stick to the matrix and not skimming through those that bind the ligand in a particular manner. This is best summarized by the dictum “you get what you select for”[136]. To select structure-switching aptamers, we need to re-think SELEX by, e.g., following the example of naturally occurring riboswitches, which possesses this capability by default[137].

1.3.1 Capture-SELEX

The methodology of Capture-SELEX was developed with precisely that in mind[138,139]. To drive the selection process towards structure-switching aptamers, this strategy uses an initial library designed to contain a constant docking region that is complementary to a capture-oligo. This oligo is immobilized onto magnetic beads and incubated with the library, thus docking the latter by Watson-Crick interactions. During the elution step, the target molecule is added to the solution and can anneal with the library sequences. However, only those sequences interacting with the target molecule by undocking from the capture-oligo will be retrieved and selected.

Capture-SELEX was initially devised to isolate structure-switching DNA aptamers to be applied as molecular beacons[138,139] but was soon applied to RNA as well[140]. The Capture-SELEX methodology demonstrates that binding-induced structural changes can be employed to drive selection pressure, and therefore pave the way to even more endearing methodologies that, perhaps one day, will allow the direct selection of RNA-based fluorescence sensors as straight-forwardly as we now do with aptamers.

1.3.2 uHTS screening platforms

In the meantime, platforms for ultra-high-throughput screening (uHTS) – such as the one applied to identify iSpinach[114] and oCoral[98] fluorogenic RNAs – can be of invaluable benefit to engineering allostery-controlled RNA sensors. Typically, the iterations of transducer modules that are tested to conjoin a ligand-binding and reporting aptamers are low: the glycine RNA sensor was an exception with circa 100 constructs screened[**127**], but the numbers usually are 5 to 10 times lower[122-126,128,129,131,132]. The low amount of screened constructs limits the effectiveness of allosteric designs. Instead, uHTS platforms allow isolating allostery-controlled RNA sensors from rationally-designed pools of up to a million different sequences.

To showcase the potential of uHTS, Autour *et al.* identified an RNA-based fluorescence sensor for theophylline employing a microfluidic platform based on *In vitro* compartmentalization (IVC) in aqueous droplets in emulsions[141]. The sensor was isolated from a library connecting the structure-switching theophylline aptamer and iSpinach via two stretches of four randomized residues, with a diversity of 65.000 sequences[130].

2 AIM OF THIS STUDY

This study aimed to develop a fluorogenic RNA sensor for fructose 1,6-bisphosphate (FBP) by selecting a variant of the constitutively fluorescent Spinach RNA that could bind its cognate dye DFHBI in an allosteric manner. For this purpose, partially-structured library designs of Spinach with riboswitch-inspired features and a tweaked sequence space were employed.

The study was based on the hypothesis that an RNA library containing the critical sequence information of Spinach fluorogenic core would increase the percentage of DFHBI-binders compared to a completely random library and thus facilitate the emergence of Spinach sensors with allosteric-binding properties.

A methodological challenge of this study was to devise an apt selection protocol that would hinge on ligand-induced structural rearrangement, thus allowing the enrichment of FBP-dependent binders from a sequence pool primed towards canonical DFHBI binders.

3 RESULTS

The experimental part of this dissertation is organized into seven sections. Section **3.1** details the steps that were put forward to establish an *in vitro* selection protocol based on affinity separation to enrich a fluorogenic RNA of DFHBI, and the selection outcome was compared to the literature of research where an analogous approach was applied. Section **3.2** examines the attempts to isolate improved variants of Spinach fluorogenic RNA, a widely characterized fluorogenic RNA of DFHBI. In section **3.4** are listed the attempts to re-purpose the affinity separation protocol to enrich ligand-responsive fluorogenic RNAs.

Section **3.5** describes the selection strategies applied to select an allosteric variant of Spinach, which responded to the metabolite Fructose 1,6-bisphosphate (FBP), and the later identification of sequence s81 via a droplet screening platform based on in vitro compartmentalization (μIVC), also analyzed in this section. Then, the binding properties of s81 towards the analyte FBP and the dye DFHBI were characterized *in vitro* and summarized in section **3.5**, where a mechanism of functioning for this sequence as an FBP sensor was laid out.

After the *in vitro* characterization, in section **3.6**, s81 was transiently transfected in different mammalian cells to identify a suitable cell line for the sensor expression,

and s81 was thereby applied to image the effect of oligomycin A, a well-known glycolytic booster, to the level of intracellular FBP. These results were further investigated in section **3.7**, where s81 fluorescence response to oligomycin A was compared head-to-head with the measured variation of intracellular FBP, and the potential of s81 as a cellular FBP sensor was re-discussed due to its loose specificity to a wide array of nucleotide analogs.

3.1 Selection of RNA aptamers that increase the fluorescence of DFHBI

To meet the goal of selecting allosteric variants of Spinach that respond to a specific analyte, the first milestone to achieve was reproducing and establishing in-house the SELEX protocol that allowed its identification. Spinach is a solid fluorogenic scaffold for this endeavor for the wealth of studies published on its characterization[94,104], optimization[83,93-95], and expression in many cellular models[78-83], which are critical to apply the identified sensor in the cellular context. Plus, Spinach was preferentially employed to other fluorogenic RNA to rationally design allosteric sensors (a complete list is found in **Table 1**), and this proofs-of concept could be of guidance for the selection.

Jaffrey and co-workers identified Spinach through a SELEX protocol based on affinity separation on Sepharose matrix, where the cognate dye DFHBI was chemically coupled through NHS chemistry[77]. The authors enriched fluorogenic sequences from a partially structured library originally designed by Szostak and co-workers[142], whereas the selection failed using a linear N70 random library. The partially structured library comprises two N26 stretches separated by a 12 nt stem-loop primed RNA pool and is depicted in **Figure 5 a**. To note, the authors reported a steady fluorescence enrichment using affinity elution from the resin with a molar excess of DFHBI from the 5th cycle up until the 10th SELEX cycle.

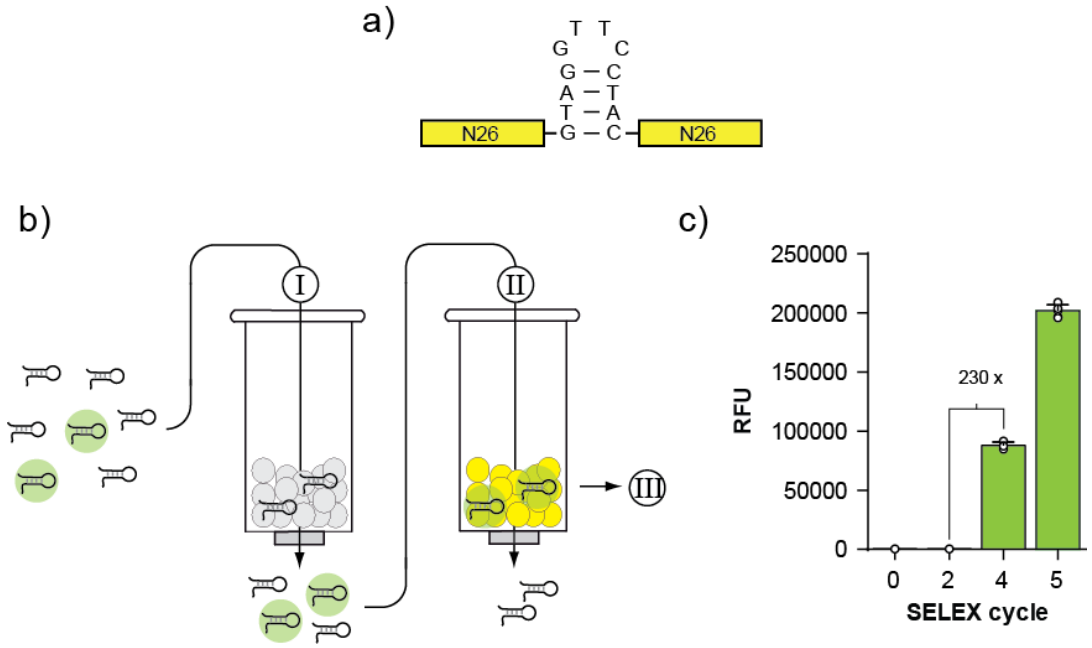


Figure 5 | Fluorogenic RNAs were selected via SELEX for DFHBI.
a, 2D representation of the 12 nt stem-loop primed library separating the random region in two stretches of 26 nucleotides, designed initially by Szostak and co-workers[142]. **b,** SELEX scheme applied for the selection of DFHBI binders. The initial random library pool is depicted as randomly structured sequences and DFHBI binders (green). The negative (**I**) and positive (**II**) selection steps are represented, and Sepharose is color-coded as mock (gray) or DFHBI-coupled (yellow). **c,** Enrichment of sequences that increase the fluorescence of DFHBI with each SELEX cycle. The fold-change increase of the fluorogenic signal from cycle two to cycle four is noted. Fluorescence assay performed at 2 µM purified RNA from the various selection rounds and 10 µM DFHBI in selection buffer (125 mM KCl and 5 mM MgCl2 in 40 mM HEPES pH 7.5).

The Sepharose-based affinity separation and the stem-loop primed RNA library design was maintained in the in-house reproduced protocol, but the amount of affinity matrix used per cycle was reduced, and a tRNA competitor was introduced to trying maximizing the stringency of selection. The affinity elution step was also substituted with a heat elution step to rely on the previous selections and experience of the lab[143,144].

	Enriched sequences aligned to the initial pool	RNA length	Frequency (%) R0	R2	R4	R5
Pool	NNNNNNNNNNNNNNNNNNNNNNNNNNGTAGTTCGCTACNNNNNNNNNNNNNNNNNNNNNNNNNNN	105				
D1	TGGTGAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCTCC	93	0.00	0.02	21.65	67.31
D2	TGGACTGAGGAGGGTAGATGTGGAGGGTAGTTCGCTACAGTGTTTACGTGCCGCACACTTGCTT	105	0.01	0.00	0.03	1.78
D3	TTGGAACGAGAGGGTGTGGAGGTGTGGTAGTTCGCTACATTCGTATGGCGTGCTGCTTGGTTGA	105	0.00	0.00	0.14	0.77
D4	TGAGAGGTGGGAGAGGAGGTTTCAGTGTAGTTCGCTACTCAATGAATGCGTCGGGATTTCCGTG	105	0.00	0.00	0.10	0.56

Table 2 | D1 is the most enriched aptamer revealed by NGS analysis on DFHBI SELEX. Sequence table of the most enriched sequences identified by NGS analysis, aligned to the structural stem-loop of the library (red). The RNA length and enrichment frequency per cycle are reported.

A schematic representation of one SELEX cycle is presented in **Figure 5 b**. Initially, the random library pool was incubated with mock Sepharose to eliminate sequences that might bind specifically to the affinity matrix; this is defined as negative selection[145,146] (step **I**). The flow-through was then incubated with DFHBI-charged Sepharose as positive selection (step **II**) and washed with selection buffer to remove non-binding sequences. Finally, the sequences bound to DFHBI-Sepharose were retrieved by heat elution (step **III**).

After five selection cycles, the purified RNA pools of the different cycles were incubated with DFHBI, and the enrichment of sequences that increase the fluorescence of the dye was tested via fluorescence assay, showcasing a strong enrichment already from the 4th section cycle (230-fold fluorescence light-up, **Figure 5 c**).

The selection cycles were analyzed by next-generation sequencing (NGS) to monitor the selection trend and identify aptamer candidates. The NGS analysis revealed that sequence D1 was predominantly enriched, with a frequency rising from cycle 2 onwards (0.02, 21.65, and 67.31 % in cycles 2, 4, 5, respectively, **Table 2**). Secondly, sequence D2 showed a 60-fold enrichment in the last selection cycle, rising from 0.03 to 1.78%, respectively. To note, sequence D1 displayed a trimming of 12 nucleotides compared to the original pool and the other top-enriched sequences (**Table 3**).

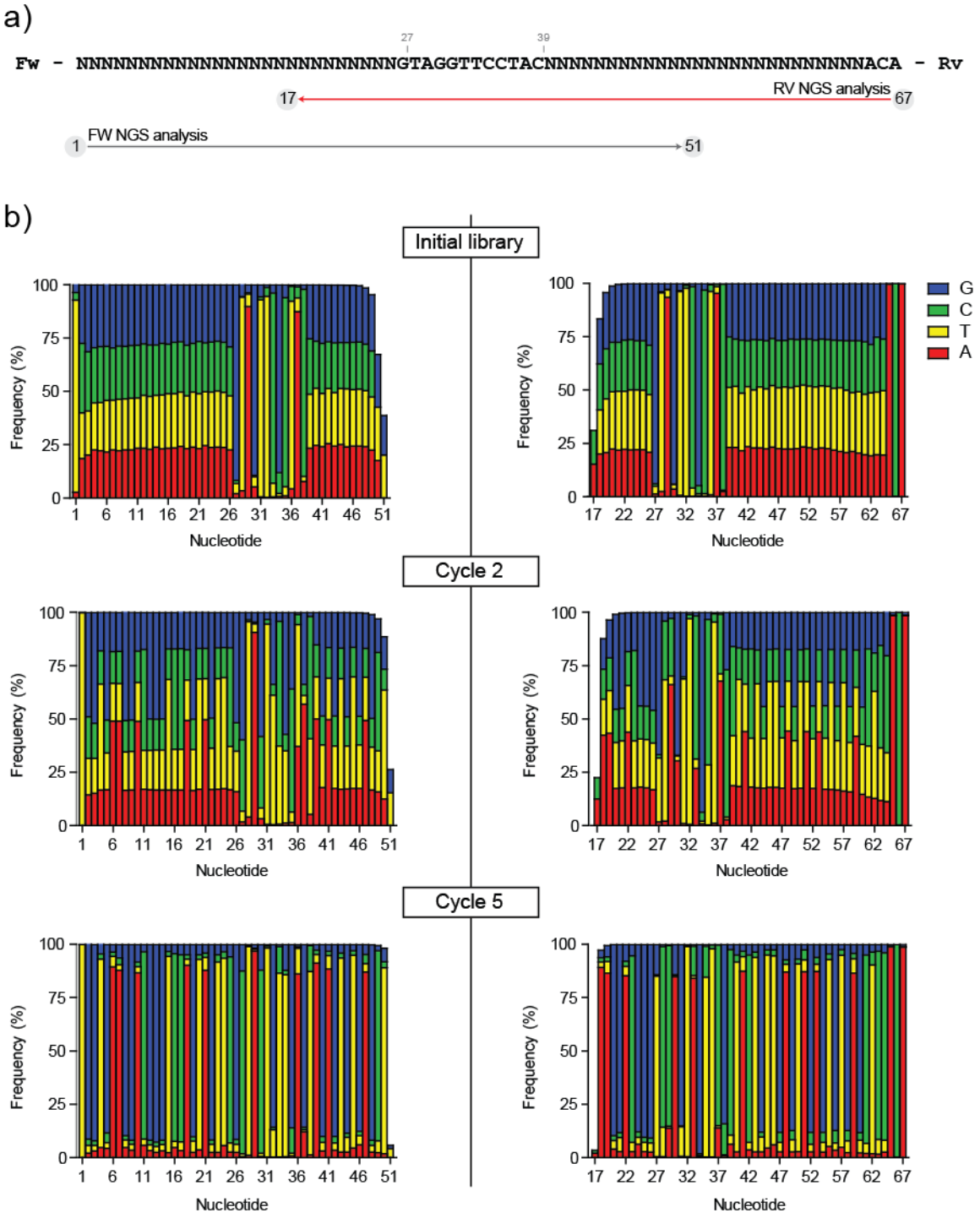


Figure 6 | DFHBI SELEX greatly enriched a unique sequence motif in NGS analysis.
a, Schematic representation of the NGS analysis performed from the forward and reverse primers of the stem-loop library.

- continues in the next page –

– continuation of previous page –

b, Unique sequence de-enrichment analysis of the initial library pool compared with cycle 2 and 5 of test DFHBI SELEX. The initial library shows a 1:1:1:1 randomization of the two 26 nt stretches on the center of which the G27-C39 stem-loop is found; this motif is constant and carried out throughout selection. Due to its 67% enrichment, aptamer D1 is the consensus sequence appearing in cycle 5 nucleotide analysis.

The enrichment of the D1 aptamer was also apparent in the nucleotide distribution analysis, which describes how every nucleotide identity of the library – random or otherwise – was affected by the SELEX pressure in each cycle (**Figure 6**). In this analysis, the initial library pool shows equal randomization of the 26N stretches beside the 12 nt stem-loop motive at the center (G27-C39 nucleotides, **Figure 6**). Instead, aptamer D1 was robustly enriched (67.31% in the 5th cycle, **Table 1**), and its sequence could be read in the consensus sequence of the nucleotide distribution of the 5th cycle.

Next, sequence D1 was compared to Baby Spinach[94] – the minimum fluorogenic motive of the Spinach aptamer – to test if the two aptamers were related. Interestingly, sequence D1 is identical to Baby Spinach by sequence composition and predictive folding (**Figure 7 a**) apart from the stretch G39-C50 of the library fixed stem-loop (**Figure 7 b**).

Next, the dissociation constant (K_D) of D1 aptamer to DFHBI was obtained indirectly by measuring the increase in fluorescence as a function of increasing fluorophore concentration in the presence of a fixed concentration of D1, assuming that every DFHBI-D1 interaction increases the dye's fluorescence. This indirect K_D measurement is commonly used to measure the affinity of fluorogenic RNAs to their cognate dye[77,98,103]. The D1 aptamer exhibited a K_D of 915 ± 57 nM (**Figure 7 d**), higher than the one reported for Spinach (537 nM[77]), and a 192-fold light-up rate with 10 µM DFHBI. In current literature, the K_D of Baby Spinach for DFHBI was

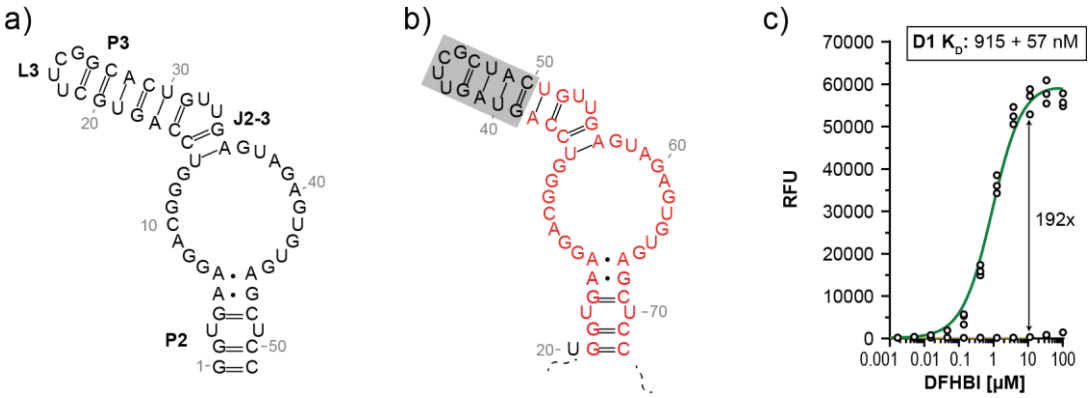


Figure 7 | D1 is homologous to Baby Spinach aptamer.
a, Secondary structural representation and nomenclature of Baby Spinach consistent with Warner KD *et al.*, where the aptamer was first documented[94]. **b,** Secondary structural representation of D1 aptamer obtained by the NUPACK web tool, which considers energy minimization[148]. The nucleotides of D1 that overlap to Baby Spinach (red), and the 12 nt constant stem-loop are highlighted (grey square). **c,** K_D measurement of 1 µM sequence D1 towards DFHBI (green) compared to DFHBI autofluorescence (yellow) in selection buffer (125 mM KCl and 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5). DFHBI is 192-fold more fluorescent in complex with D1 aptamer at 10 µM dye concentration.

never measured or disclosed in all papers which made use of this aptamer[94,115,117,118,147].

The results indicated that sequence D1 corresponded to Baby Spinach by sequence comparison and induction of DFHBI fluorescence (**Figure 7**) and that the SELEX protocol on DFHBI was successfully reproduced from the published results of Jaffrey and co-workers[77].

3.2 Selection of Spinach aptamer improved variants

The Spinach-DFHBI fluorogenic system was widely applied as a fluorogenic RNA-dye system, notwithstanding its photochemistry and ion-requirements are not optimal for microscopy settings[83] and expression in the cytoplasmic milieu[95], respectively. Given that the selection of Spinach aptamer was reproduced efficiently and with high discrimination (**Figure 5, Figure 6**), the selection of improved variants of the Spinach-DFHBI pair was investigated.

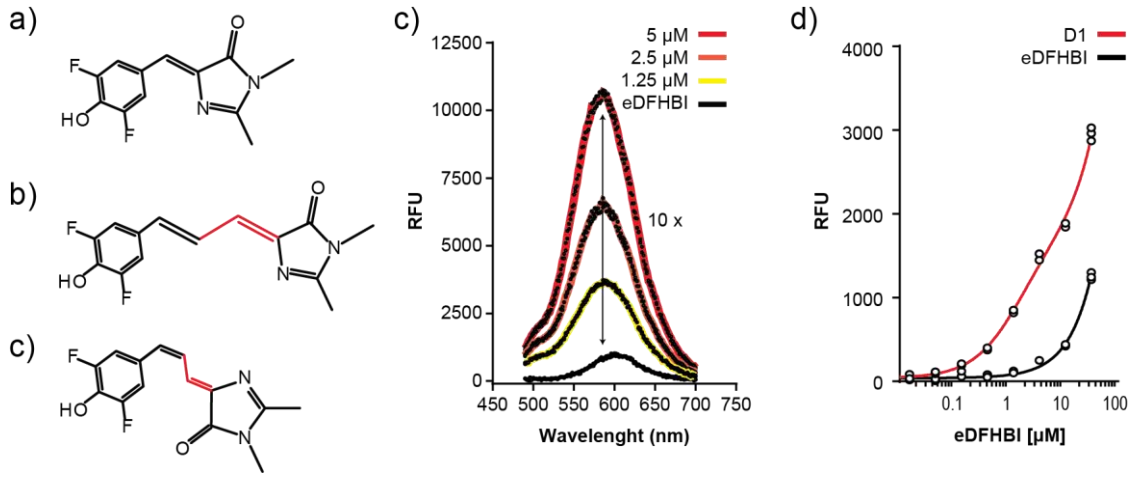


Figure 8 | eDFHBI-mix exhibits background interaction with Baby Spinach.
a-b, Chemical structure of DFHBI (**a**) compared with (E)-eDFHBI (**b**) and (Z)-eDFHBI (**c**). The vinylene extension of the methine bridge is highlighted (red). **c,** Increase of eDFHBI fluorescence, showing an emission maximum a 600 nM, upon incubation with increasing concentration of D1 aptamer, reaching a 10-fold increase at 5 µM RNA. **d,** Dose-response fluorescence interaction showcasing the autofluorescence curve of the eDFHBI-mix (black) compared to the fluorescence with 1 µM D1 (red) in selection buffer (125 mM KCl and 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5).

First, the critical aspects of Spinach-DFHBI were tackled by fine-tuning the SELEX protocol to identify variants of the Spinach aptamer that fluoresce in the red-spectrum. Red-emitting fluorescent dyes are better suited for imagining living cells and tissues that green-fluorescent ones, for they exhibit reduced autofluorescence from the biological matrix, permit deeper tissue penetration, and avoid visible-light absorption[149].

To that end, Dr. Rosenthal synthesized a variant of DFHBI, named extended DFHBI (eDFHBI), in which the methine bridge connecting the 2,6-difluorophenole and imidazolinone rings was extended by one vinylene unit (**Figure 8 a- c**)[150]. This extension strategy is well-described in the literature to obtain dyes exhibiting a red-shifted emission wavelength and broader Stoke shifts, e. g. the Cy3 to Cy5 cyanine synthesis[151] and the two different derivatives of thiazole orange TO1 and TO3, for which Mango I-III fluorogenic RNA series was developed[100-102]. However, the

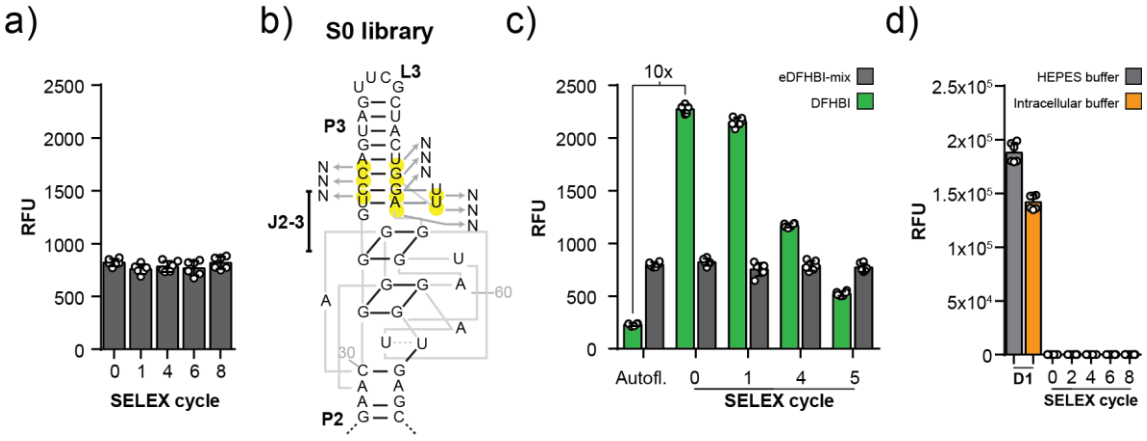


Figure 9 | SELEX did not enrich improved variants of Baby Spinach.
a, No sequences that increase eDFHBI-mix fluorescence were enriched from the stem-loop primed RNA library after 8 SELEX cycles. **b,** The S0 library comprises the stem P2, G-quadruplex core, and the stem-loop PL3 of Baby Spinach, whereas the junction J2-3 and part of stem P3 (yellow) were randomized to select an aptamer with a larger pocket that might better accommodate one eDFHBI isomer. The structure of Baby Spinach was obtained from Warner *et al.*,[94]. **c,** No sequences that increase eDFHBI-mix fluorescence were enriched from the S0 library after 5 SELEX cycles (grey bars). Also, the selection de-enriches the DFHBI-binders that are present in the initial library pool, as shown by the 10-fold light-up change between DFHBI autofluorescence and cycle 0 (green bars). **d,** No sequences that increase DFHBI fluorescence were enriched from the stem-loop primed RNA library after 8 SELEX cycles in the intracellular mimicking buffer (160 mM KCl, 5mM NaCl, 1 mM $MgCl_2$ in Phosphate buffer pH 7.2) albeit aptamer D1 exhibits the comparable fluorogenic properties for DFHBI in both the intracellular (orange) and HEPES buffers (grey). Fluorescence assay performed at 2 µM purified RNA from the various selection rounds or D1 aptamer and 10 µM dye. Unless specified, assays were conducted in selection buffer (125 mM KCl, 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5).

synthesis was not stereospecific, and a mixture of two isomers (E)-eDFHBI (**Figure 8 b**) and (Z)-eDFHBI (**Figure 8 c**) was obtained[150].

The photometric properties of the eDFHBI mixture were evaluated in combination with the D1 fluorogenic RNA to test if one or both isomers in the mixture retained some interaction with this aptamer (**Figure 8**).

The eDFHBI mixture exhibited red-shifted emission (λ_{em}= 600 nM, **Figure 8 c**) and an optimized Stoke shift ($\Delta\lambda$=145nm, **Figure 8 c**) compared to DFHBI (λ_{em}= 505 nM, $\Delta\lambda$=45 nM[77]). Furthermore, eDFHBI showed an increment of the emission

fluorescence when incubated with increasing concentration of D1 aptamer, but only at high RNA levels (1 µM onwards, **Figure 8 b**). To compare, 10 µM DFHBI interacted linearly with D1 between 0.2 nM – 5000 nM, whereas 10 µM DFHBI-mixture required >300 nM of D1 to obtain an increase in fluorescence, as shown in the dose-response fluorescence assay in **Supporting Figure 1**. The dissociation constant of D1 to eDFHBI-mixture could not be measured, given that D1 exhibited only a background fluorescence increase to the eDFHBI-mix (red curve, **Figure 8 c**) compared to eDFHBI-mix autofluorescence (black curve, **Figure 8 c**).

Despite the low affinity with D1, the aptamer recognized the eDFHBI-mix hinted that it might be possible to select a more affine variant within the stem-loop primed RNA library (**Figure 5 a**), from which D1 was selected. Eight SELEX cycles were performed on a Sepharose matrix functionalized with the eDFHBI-mix, and the selection cycles were then tested by fluorescence assay, but no intensification of the dye's fluorescence was found (**Figure 9 a**).

To optimize the initial library towards a sequence space that closely resembled Baby Spinach, the Spinach 0 (S0) library was designed containing the core structural information of the aptamer. S0 library consisted of all the critical folding units of Baby Spinach – stem P2, stem-loop PL3, and G-quartet – except junction 2-3 (J2-3, **Figure 9 b**). The junction J2-3 interacts with the distal imidazolinone ring of DFHBI (**Figure 3**) and engulfs it inside a pocket obtained with the G-quadruplex domain[77]. The J2-3 unit, including part of stem P3, was randomized to select an optimized variant of Baby Spinach with a larger pocket that might better accommodate eDFHBI, which is sterically bulkier than DFHBI (**Figure 8 a, b**). With this design rationale, the obtained S0 library contained nine randomized nucleotides and a sequence space of 262144 variants (**Figure 9 b**).

Due to its similarity to Baby Spinach, the S0 library was tested for background fluorescence for both DFHBI and eDFHBI-mix (**Figure 9 c**). The S0 library did not increase eDFHBI-mix fluorescence, whereas it showcased a 10-fold fluorescence

increase with DFHBI. This background affinity to DFHBI indicated that few variants of the library retained the paternal J2-3 unit, which was expected considering the randomization targeted only nine nucleotides of Spinach.

Five SELEX cycles were performed using the S0 library to enrich fluorogenic sequences of the eDFHBI-mix (grey bars, **Figure 9 c**), while the DFHBI background fluorescence was also tested in parallel (green bars, **Figure 9 c**). Throughout the selection, no fluorogenic binders of eDFHBI-mix were enriched, and concomitantly, the background binders of DFHBI were de-enriched, as shown by the decrease of DFHBI light-up change from cycle 0 onwards (green bars, **Figure 9 c**).

The decrease in DFHBI background binders, which are likely similar to D1, might indicate that while D1 showed minimal interaction with the eDFHBI-mix (**Figure 8**), it was not sufficiently affine to be enriched within SELEX. Therefore, the targeted randomization starting from D1 was not suitable to identify fluorogenic RNAs of eDFHBI. The presence of two stereoisomers in the eDFHBI-mix (**Figure 8 b,c**) might also explain why SELEX on this compound could not succeed.

Next, the selection of optimized aptamers for DFHBI that fold correctly in an intracellular-mimicking buffer system was investigated. Paige *et al.* selected Spinach in 125 mM KCl, 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5[77], same as the one used to identify D1 aptamer (**Figure 5**). This buffer was used in many SELEX experiments[152-154] but is not designed to mimic the intracellular milieu[77]. To mimic the physiological conditions of the cytoplasm[155], a phosphate buffer rich in K^+ and decreased Mg^{2+} to 1 mM was employed. Baby Spinach shows comparable light-up properties in this buffer system and the previously used HEPES-based buffer (grey vs. orange bars for D1 control, **Figure 9 d**). Nevertheless, this SELEX protocol did not enrich fluorogenic sequences (**Figure 9 d**).

In sum, both attempts to *in vitro* select a red-shifted variant of Spinach and a version of this RNA that requires low Mg^{2+} were unsuccessful (**Figure 9**). However, the incapability of re-selecting D1 or similar aptamers in the intracellular-mimicking

buffer was baffling. To note, Jaffrey and co-workers succeeded in selecting Corn fluorogenic RNA using DFHBO, a second-generation DFHBI dye, in a HEPES-based buffer containing 1 mM $MgCl_2$[97]. This could indicate that a phosphate-based buffer might not be suitable for the selection of this class of fluorogenic RNAs, and the fact that phosphate groups weakly complex Mg^{2+}-ions strengthens this hypothesis[156].

3.3 Selection of ligand-dependent allosteric binders of DFHBI via SELEX

3.3.1 Allosteric SELEX on the stem-loop primed RNA library

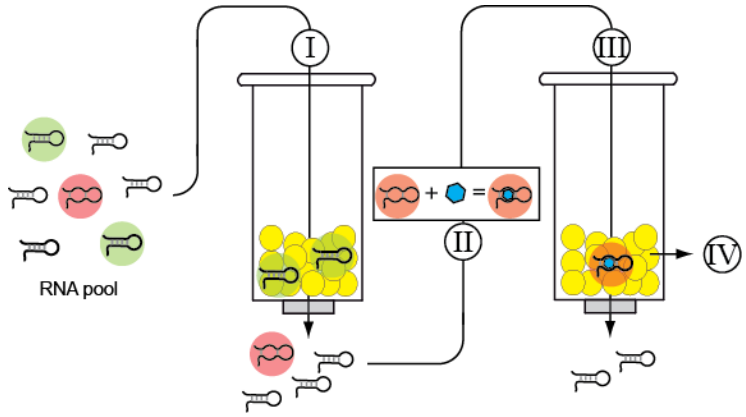


Figure 10 | Allosteric SELEX protocol to select target-dependent fluorogenic RNAs. SELEX scheme applied for the selection of allosteric binders of DFHBI. The initial library RNA pool (left) representatively depicts three different sequence populations: randomly structured sequences (black), DFHBI binders (green), and target-dependent DFHBI allosteric binders (red). To de-enrich the DFHBI binders, the library is initially incubated with DFHBI sepharose as negative selection step (**I**); the flow-through RNA is incubated with the target (blue hexagon) to allow the structure-rearrangement red sequences, whereas the remaining sequence will not interact with the metabolite (**II**). The solution is incubated with DFHBI Sepharose as a positive selection step (**III**). Finally, the binding sequences are washed with selection buffer plus FBP to remove unspecific binders, and finally, the red sequences are eluted by heat elution (**IV**).

The selection of a Spinach-homologous fluorogenic RNA, named D1, indicated that Jaffrey and co-workers' published results could be reproduced and that SELEX based on affinity separation with DFHBI-Sepharose was a robust protocol to which to build up more complex SELEX schemes (Section 3.1).

Next, the affinity separation protocol was combined with the capability of RNA to structurally rearrange itself[125], with the goal of selecting target-dependent fluorogenic RNAs of DFHBI.

To that end, the allosteric SELEX protocol was streamlined to give a selective advantage to sequences that exhibit a target-dependent binding to DFHBI over canonical DFHBI binders. A schematic representation of one allosteric SELEX cycle is presented in **Figure 10** and comprises three subsequent steps. First, the RNA library was incubated with the DFHBI-charged matrix (step I, **Figure 10**) as a negative step to deplete canonical DFHBI-binders. Once the canonical binders of DFHBI were de-enriched, the structural rearrangement of allosteric binders of DFHBI was promoted by incubating the flow-through RNA with the ligand (step II, **Figure 10**). The ligand-dependent binders are now capable of binding DFHBI, and

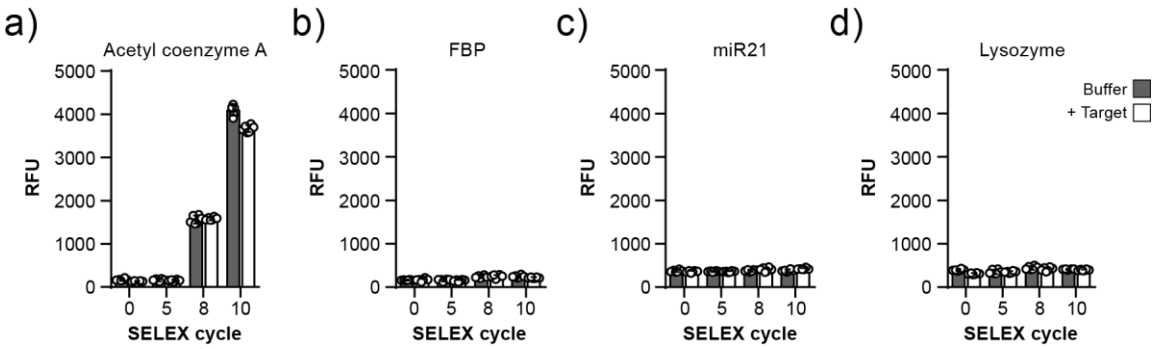


Figure 11 | Allosteric SELEX on the stem-loop library failed to enrich target-dependent light-up sequences.
a-d, Fluorescence binding assays of the purified RNA from the different SELEX cycles tested in the absence (grey bars) or presence (white bars) of AcCoA (200 µM, **a**), FBP (5 mM, **b**), miR21 (1 µM, **c**) and lysozyme (0.2 µM, **d**). Fluorescence assay performed at 1 µM purified RNA from the various selection rounds and 10 µM DFHBI in selection buffer (125 mM KCl, 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5) using the target concentration applied in the selection.

they can be retrieved in a positive incubation step with DFHBI-charged matrix (step III, **Figure 10**) and consequent washing and heat elution (step IV, **Figure 10**).

The capability of the designed allosteric SELEX protocol to enrich ligand-dependent binding to DFHBI from stem-loop primed library (**Figure 5 a**) was tested on fructose 1-6 bisphosphate (FBP) and a panel of a small molecule (acetyl coenzyme A), a microRNA (miR21), and a protein target (lysozyme). These targets were chosen arbitrarily for their widely different chemical nature to screen different functional groups that could interact with RNA and maximize success. While the level of FBP in bacteria and mammal cells was quantified to be in the millimolar range[157], acetyl coenzyme A ranges the micromolar[158]. Thus, their concentration in SELEX was set to 5 mM and 200 µM, respectively. Instead, the SELEX concentration of the ligands miR21 (1 µM) and lysozyme (0.2 µM) was set by fluorescence quenching of D1 (**Supporting Figure 2**).

To test the selection, the enrichment of sensor sequences was indirectly measured by comparing the background fluorescence of the pools from the subsequent SELEX cycles with (+ Target, **Figure 11**) and without the target (Buffer, **Figure 11**) to identify those selections whose fluorescence would markedly increase when the target was present. The fluorescence assay for FBP (**Figure 11 b**), miR21 (**Figure 11 c**), and lysozyme (**Figure 11 d**) revealed that no fluorogenic sequences were enriched throughout the selection. On the contrary, the selection for acetyl coenzyme A enriched for DFHBI fluorescence, but not target-dependent (**Figure 11 a**).

While no evidence was obtained that sensor sequences were enriched, the null background fluorescence for FBP, miR21, and lysozyme (**Figure 11 b-d**) revealed that the percentage of canonical DFHBI binders could be maintained low by the negative selection step of the allosteric SELEX protocol (step II, **Figure 10**). However, the negative selection step was not effective with acetyl coenzyme A

(**Figure 11 a**), indicating that also the target played a role in the negative selection by, most likely, further stabilizing or de-stabilizing DFHBI-binders.

Based on these results, the possibility to enrich sensor sequences via SELEX was not yet dismissed. Instead, it was concluded that the sequence space comprised in the stem-loop primed library was not suited for identifying allosteric binders of DFHBI. It stands to reason that the percentile of the sequence population capable of binding DFHBI in an allosteric fashion is exponentially smaller than the canonical binders of DFHBI. These are already insignificant inside the sequence space of the stem-loop library, as shown by the NGS analysis of the test-selection for DFHBI that exhibited a null relative frequency of D1 aptamer in the initial stem-loop pool (**Table 2**). Thus, the library had to be re-designed to comprise a sequence space that better suited the purpose of the selection.

3.3.2 Partially-structured libraries of Baby Spinach

To increase the allosteric population in the library, specialized library designs containing the structural information of Baby Spinach were investigated (**Figure 12**). From a structural point of view, it can be reasoned that allosteric binder might reflect the Baby Spinach sequence, given that they both need to bind DFHBI. Therefore, the chance of selecting a variant that fluoresces target-dependently may be increased when equipping the library with the module of Baby Spinach.

With that in mind, the crystallographic information available for the Spinach-DFHBI pair (Section 1.2.2.1) was used to design a panel of libraries in which the critical stems that support the G-quadruplex domain of Baby Spinach were progressively randomized from the P3 stem (**Figure 12 a-c**) and P2 stem sides (**Figure 12 e-g**). From the P3 stem-side of Baby Spinach, library S1 lacked the distal part of the stem, which was replaced by two N20 regions (**Figure 12 a**). The full stem and part of the junction 2-3 – critical for DFHBI interaction – was randomized in the library S2 (**Figure 12 b**). Finally, library S3 was designed by randomizing the whole region above the G-quadruplex, which was replaced by an N40 (**Figure 12 c**).

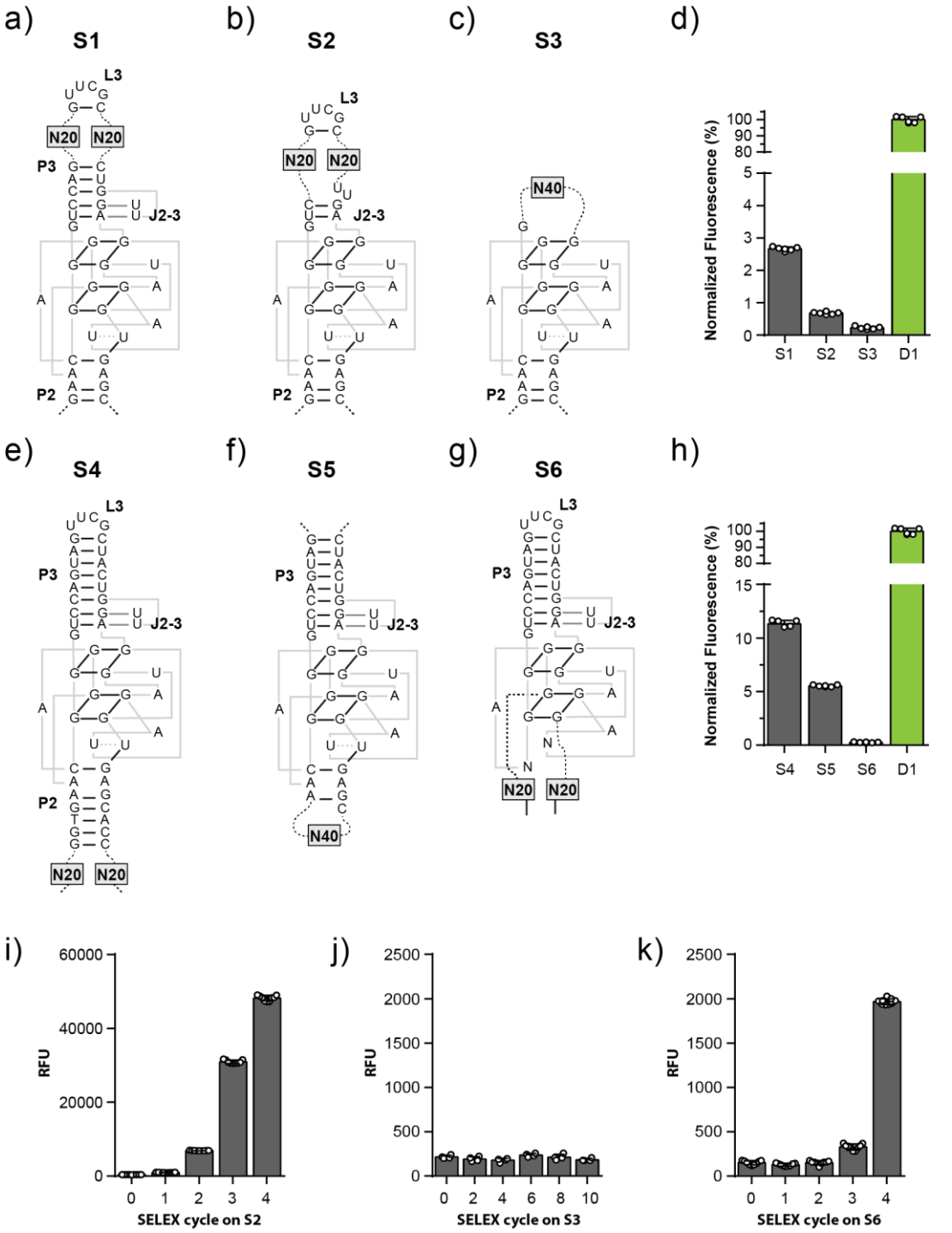


Figure 12 | S2 and S6 were picked as partially-structured libraries of Baby Spinach. **a-g,** Design of structurally-driven libraries of Baby Spinach where the critical domains which support the folding of the core G-quartet were progressively randomized from the stems P3 (**a-c**) and P2 (**e-g**) sides.

- continues in the next page –

– continuation of the previous page –

S1 (**a**) lacks the distal P3 stem, whereas S2 (**b**) the full stem and part of J2-3, and S3 (**c**) the whole region above the G-quartet. S4 (**e**) lacks the distal P2 stem, S5 (**f**) the full stem, and S6 (**g**) the whole region below the G-quartet. **d, h,** Background fluorescence analysis of the sequences that increase the fluorescence of DFHBI within the libraries S1 to S3 (**d**) and S4 to S6(**h**), normalized to D1 aptamer control, which was set to 100%. **i- k,** DFHBI SELEX test applied on library S2 (**i**), S3 (**j**), and S6 (**k**), which showed <1% fluorescence of D1 aptamer in the background binding analysis. Fluorescence assay performed at 1 µM purified RNA from the various selection rounds or D1 aptamer and 10 µM DFHBI. Assays were conducted in selection buffer (125 mM KCl, 5 mM MgCl2 in 40 mM HEPES pH 7.5).

The P2 stem-side of Baby Spinach was randomized in the same manner: library S4 was designed by randomization of the distal part of P2, and S5 lacked the full stem (**Figure 12 d, e**). Finally, S6 was designed by randomizing the whole region below the G-quadruplex, which was replaced by two N20 regions (**Figure 12 f**).

Next, the background fluorescence of the different libraries was measured to estimate which library design would more strongly de-stabilize the Baby Spinach module by comparing it to D1 fluorescence (**Figure 12 d, h**). The background fluorescence of the libraries correlated to the extent of randomization of Baby Spinach: the highest fluorescence was obtained with S1 (11% of D1 fluorescence, **Figure 12 d**) and S4 (2.7% of D1 fluorescence, **Figure 12 h**), where Baby Spinach was less randomized. Then, the fluorescence progressively dropped in S2 and S5 (5 and <1%, **Figure 12 d, h**) and was minimal in S3 and S6 (both <1%, **Figure 12 d, h**).

Those libraries with a background fluorescence >1% were excluded with the assumption that it would be unlikely to identify allosteric binders in a library so markedly primed by canonical DFHBI binders. Then, the three remaining libraries – S2, S3, and S6 – were tested in a DFHBI SELEX protocol (**Figure 5 a**) to anticipate their propensity in enriching canonical DFHBI binders. These selections were compared to our benchmark selection with the stem-loop primed library (**Figure 5 c**).

Based on the fluorescence increase throughout the selection, only S2 performed as expected, showing a marked fluorescence increase already at cycle 2 (**Figure 12 i**), whereas no fluorescence enrichment was shown in the same cycle of the stem-loop library (**Figure 5 c**). The increase of fluorescence to an earlier selection cycle is an indication that this library is more prone to enrich for DFHBI binders. Instead, S3 showed comparable fluorescence enrichment to the stem-loop library with marked fluorescence enrichment at cycle 4 (**Figure 12 k**).

It was surprising to find that S6 showed no increase of fluorescence even after 10 cycles (**Figure 12 j**), indicating that this library was a worse pool for selecting DFHBI-binders than the stem-loop library (**Figure 5 c**). A net loss in the capability of S6 to re-select DFHBI binders contradicted the assumption that is adding structural information of a specific aptamer sequence has the ending effect of promoting the reselection of that specific aptamer motif. However, given the relatively small number of tested library designs and the fact that these selections were not performed in parallel, this evidence is not conclusive.

Despite this inconsistency, P2 and P6 were shortlisted as partially structured libraries of Baby Spinach for their low background fluorescence (**Figure 12 d, h**) and capability of re-selecting DFHBI-binders (**Figure 12 i, k**).

3.3.3 Allosteric SELEX on partially-structured libraries of Baby Spinach

The shortlisted libraries P2 (**Figure 12 b**) and P6 (**Figure 12 g**) were tested in the allosteric SELEX protocol introduced in **Figure 10** for selecting ligand-dependent DFHBI binders that respond to the second messenger adenosine 3′, 5′-cyclic monophosphate (cAMP). Advantageously, cAMP is a ribonucleotide-derived small molecule that is a class of ligands that have continuously been used and researched for being “aptamerogenic”[142,154,159], mainly due to their capability to interact with aptamers by Watson-Crick base pairing. Apropos, Nutiu and co-workers based their innovative structure-switching selection protocol appositely on

a large panel of these ribonucleotide compounds[160], which were set to 1 mM during selection. Thus, the concentration of cAMP in the allosteric SELEX protocol was also set to 1 mM.

At first, the selections for both P2 (**Figure 13 a**) and P6 (**Figure 13 c**) enriched canonical DFHBI binders, as shown by the increase of cAMP-*independent* background fluorescence in the enrichment analysis of the different pools. This enrichment indicated that the negative selection step had to be more stringent for these ad hoc designed pools than the stem-loop library (**Figure 11**). Thus, the negative selection was iterated for three consecutive steps, instead of only one, to curb the enrichment of DFHBI-binders.

Due to the increase in negative pressure, both selections from libraries P2 (cycle 1 to 8, **Figure 13 b**) and P6 (cycle 4 to 10, **Figure 13 d**) exhibited steady background fluorescence, and consequently, a constant frequency of DFHBI-binders. Although the enrichment of DFHBI-binders was curbed, the selection with P2 did not show cAMP-dependent fluorescence increase (**Figure 13 b**). Instead, a significant increase of the cAMP-dependent fluorescence was noted in cycle 11 of the P6 library selection (p-value: 0.002, **Figure 13 d**), where the negative selection step was tentatively decreased back to 1 step.

To increase this low but significant response in the subsequent cycles of selection, an enrichment test for cycle 12 was performed at increasing negative step – 5 steps in cycle 12.1, 3 steps in cycle 12.2, and 1 step in cycle 12.3 – with the assumption that one of these conditions might be more suitable for increasing cAMP-dependent binders of DFHBI. Cycles 12.1-3 were compared to cycle 11 for reference, which showed a lower but significant cAMP-dependent fluorescence (p-value: 0.0213, **Figure 13 e**). While the background fluorescence of the library (grey bars, **Figure 13 e**) progressively diminished accordingly to the negative selection steps, the cAMP-dependent fluorescence response was lost (white bars, **Figure 13 e**).

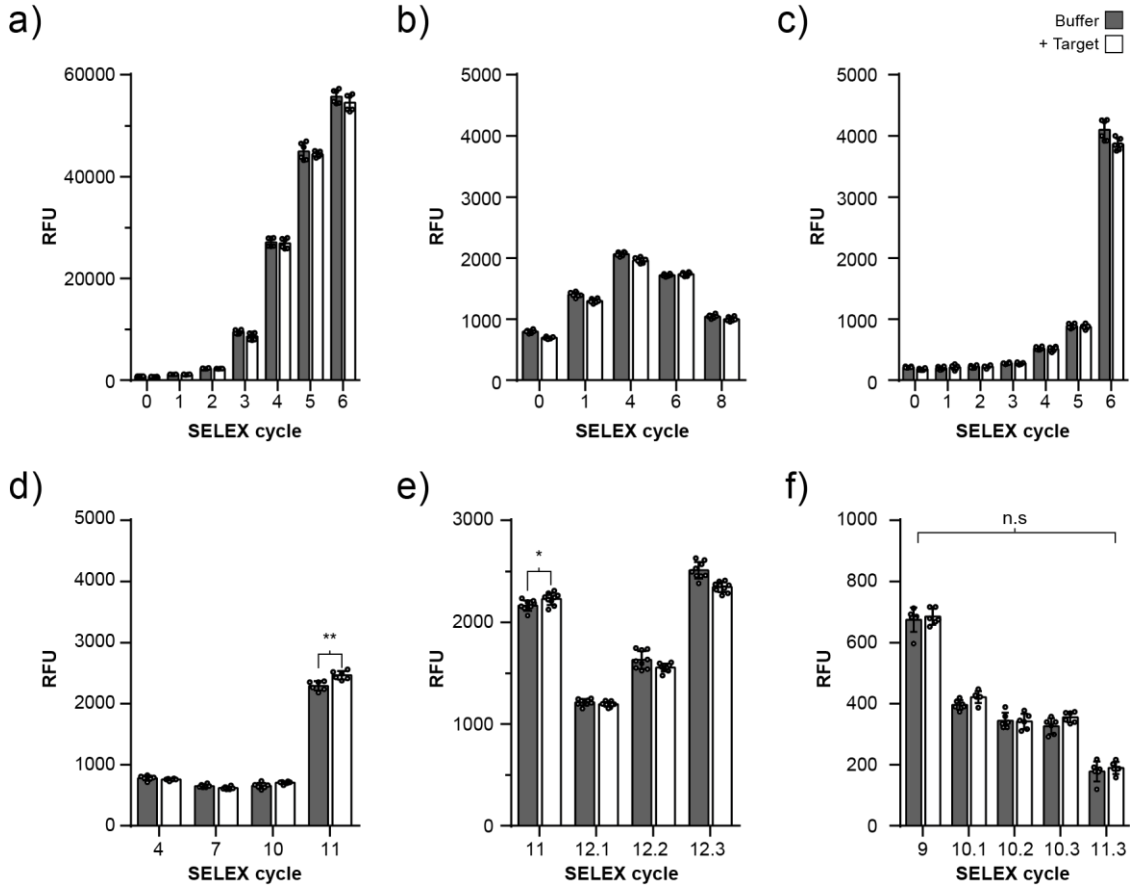


Figure 13 | Allosteric SELEX failed to enrich cAMP-responsive fluorogenic RNAs from S2 and S6 libraries.

a-f, Fluorescence binding assays of the purified RNA from the different SELEX cycles using S2 (**a-b**) and S6 libraries (**c-f**) tested in the absence (grey bars) or presence (white bars) of 1 mM cAMP. **a,** Allosteric SELEX on S2 using 1x negative selection step. **b,** Allosteric SELEX on S2 using 3x negative selection steps. **c,** Allosteric SELEX on S6 using 1x negative selection step. **d,** Allosteric SELEX on S6 using 3x negative step started from cycle 4 of SELEX c, whereas cycle 11,3 (p-value:0.002) was kept at 1x negative selection. **e,** Allosteric SELEX test using cycle 11 (p-value:0.021) of SELEX d with progressively decreasing negative selections steps: 5x (12.1), 3x (12.2), and 1x (12.3). **f,** Allosteric SELEX test using cycle 9 of SELEX d to cycle 10, where the negative selection step was progressively increased from 3x (10.1) to 6x (10.2) to 12x (10.3). Cycle 10.3 was used for another cycle of selection at 12x negative selection steps to obtain cycle 11.3. Fluorescence assay performed at 1 µM purified RNA from the various selection rounds or D1 aptamer and 10 µM DFHBI. Assays were conducted in selection buffer (125 mM KCl, 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5). P-values were calculated using the Mann-Whitney non-parametric test

As a last attempt to reconstitute the cAMP-dependent fluorescence obtained in cycle 11, the selection with library P6 was restarted from cycle nine with increasing negative selection steps – 3 steps in cycle 10.1, 6 steps in cycle 10.2, and 12 steps in cycle 10.3– and the cycles were compared to cycle 9 for reference (**Figure 13 f**). Interestingly, the increase in the negative selection steps did not affect the background fluorescence of 10.1-3, likely indicating that canonic DFHBI binders could not be de-enriched further.

However, no significant cAMP-dependent fluorescence was noted, even with a subsequent cycle at 12 negative selection steps (cycle 11.3, **Figure 13 f**).

While no cAMP-dependent fluorescence was consistently enriched, library P6 showed interesting characteristics in response to the allosteric SELEX protocol, from its sensitivity to negative selection (**Figure 13 c-d**) to some indication of allosteric fluorescence (**,**Figure 13 c**). Therefore, library P6 was tested with the panel of targets previously employed with the stem-loop library (**Figure 11**). Learning from the previous selection with cAMP, the negative selection steps were gradually increased throughout the selection to avoid a too stringent negative

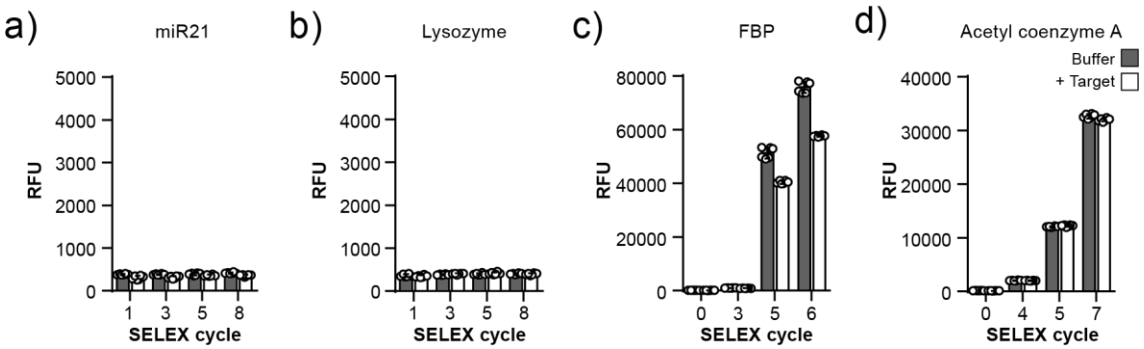


Figure 14 | Allosteric SELEX on P6 library for a panel of target ligands
a-d, Fluorescence binding assays from the different SELEX cycles on library P6 tested in the absence (grey) or presence (white) of miR21 (1 µM, **a**) and lysozyme (0.2 µM, **b**), FBP (5 mM, **c**) and AcCoA (200 µM, **d**). Fluorescence assay performed at 1 µM purified RNA from the various selection rounds and 10 µM DFHBI in selection buffer (125 mM KCl, 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5) using the target concentration applied in the selection.

pressure in the early selection cycles: Thus, the negative selection was repeated once in cycle 1, twice in cycle 1, twice in cycle 2-3, and thrice from 3-onwards. However, this optimized allosteric SELEX protocol also failed to enrich target-dependent fluorogenic sequences using miR21 (**Figure 14 a**), lysozyme (**Figure 14 b**), Fructose 1,6-bisphosphate (FBP, **Figure 14 c**) and Acetyl-coenzyme A (AcCoA, **Figure 14 d**).

Altogether, the selection of allosteric binders from either randomized (**Figure 11**) or partially randomized RNA pools (**Figure 13, Figure 14**) did not succeed. Thus, the next step was to test even less random pools comprising the information of a target-binding aptamer.

3.4 Selection of an FBP dependent light-up RNA aptamer of DFHBI

3.4.1 Design of a chimeric library of Baby Spinach and C45 aptamers

Since a generalized protocol to enrich allosteric binders of DFHBI for any given ligand did not succeed, the following efforts were focused on the identification of an allosteric fluorescence sensor for Fructose 1,6-bisphosphate (FBP), which is a key-metabolite of glycolysis and a possible proxy for glycolytic flux in different organisms (section 1.1.4). Alongside its biological importance, FBP was chosen as a suitable target due to the availability of C45[161], a specific RNA aptamer previously selected by SELEX to bind FBP.

C45 is a 51-nucleotide RNA aptamer selected by a conventional SELEX scheme, where FBP was immobilized on Epoxy-Sepharose, and 13 cycles of selection were performed starting from an N40 library[161]. Ortega *et al.* characterized the binding affinity to be specific for FBP in the low millimolar range (K_D: 6.7 mM[162]). Further, they identified the secondary structure of C45 aptamer with the structural changes occurring upon FBP binding by secondary structure prediction coupled with

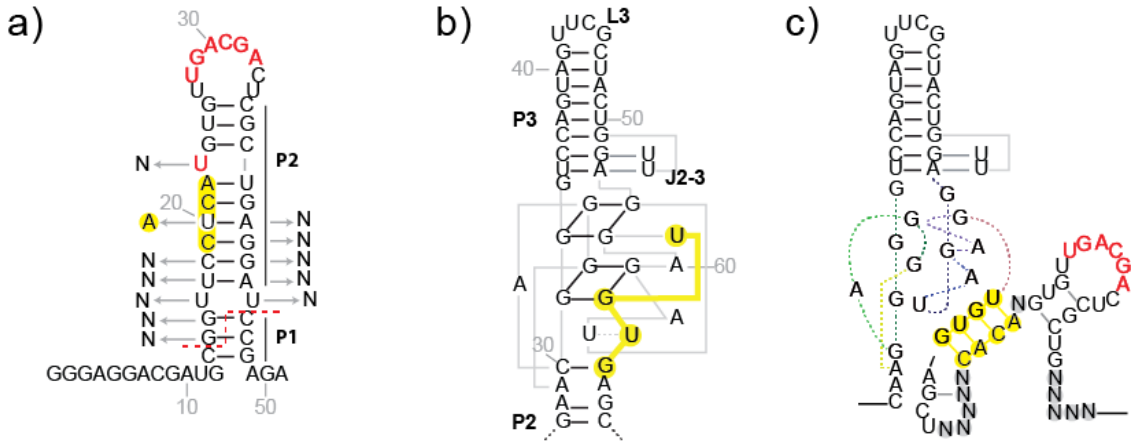


Figure 15 | Design of a chimeric library where Baby Spinach and the critical nucleotides of C45 are conjoined.
a, Secondary structure of the FBP-binding RNA aptamer C45 according to Ortega *et al.*[162]. The nucleotides showing a variation in the reactivity pattern towards 1-methyl-6-nitroisatoic anhydride (1M6) in SHAPE analysis are shown in red. U20 in P2 was replaced by an adenine nucleotide to generate a stretch of four nucleotides (nts 19-22, highlighted in yellow) complementary to the G-quadruplex of Baby Spinach (nts 62-65, highlighted in yellow, **b**). C45 aptamer was miniaturized (red dotted line), and U23 was randomized together with 10 nucleotides from the base of stems P1 and P2. **c,** Secondary structure prediction of the chimeric library depicting the G-quartet folding inhibition by the interaction between the CACA stretch of C45 and the UGUG stretch of Baby Spinach. The nucleotides belonging to the transducer (yellow), the random (grey), and FBP-interacting regions (red) are highlighted.

selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) chemical probing[162].

C45 structural predictions and FBP-dependent variations in the reactivity pattern are summarized in **Figure 15 a**. The aptamer folds into a stem-loop structure with a 12-nt stem (C13-G26 and C36-G48, **Figure 15 a**) and a 9-nt distal loop (**L**, U27-U35), that contains a 6-nt stretch most reactive to FBP by SHAPE analysis. The stem splits in a shorter basal stem (**P1**) and an apical stem (**P2**) by a uridine mismatch (U16, U45); P2 also contains a base-bulge (U23) that is reactive to FBP but, upon further characterization, was found to be inessential for FBP interaction.

Based on these data, a chimeric library was designed by conjoining the critical nucleotides of C45 for FBP interaction to the fluorogenic scaffold of Baby Spinach

(**Figure 15 b**). The previous strategies for designing allostery-regulated RNA elements were used to guide the assembly of the two aptamers (section 1.2.3). Since C45 interacts with FBP from its loop region, a *sequence-sequestering* can be used to combine the DFHBI-binding domain of Baby Spinach, following the example of the conversion from the natural TPP-sensing riboswitch into a TPP fluorogenic biosensor[125].

Initially, nucleotide U20 in the P2 stem was exchanged for an adenine residue, thereby generating a four nucleotide stretch (residues 19-22, CACA) that is complementary to the residues 62-65 (UGUG) of Baby Spinach's G-quadruplex (yellow-highlighted nucleotides, **Figure 15 a, b**). This anti-complementarity was designed to promote, in the absence of FBP, the interaction of the CACA stretch of C45 with G-quartet formation of Baby Spinach, hence impairing its interaction with DFHBI. Then, the 3'- of Baby Spinach and 5'-termini of C45 were conjoined to overlap the CACA-stretch of C45 to the UGUG-stretch of Baby Spinach (**Figure 15 c**).

Finally, the nucleotides of stem P1 and part of P2 were randomized (**Figure 15 a**), thereby creating a random region within the chimeric library to select the best allosteric design (**Figure 15 c**). This chimeric library has 11 random nucleotides, covering a sequence space of ~ $4*10^6$ variants.

Thus designed, the chimeric library was subjected to two different methodologies for selecting FBP-responsive fluorogenic sequences in parallel. The first was the allosteric SELEX protocol, initially designed (**Figure 10**) and fine-tuned in the negative selection step throughout the previous selections (**Figure 13, Figure 14**). The second methodology was based on microfluidic-assisted *in vitro* Compartmentalization (µIVC) and droplet screening[114,163]. To note, the sequence diversity of the chimeric library of circa four million variants was specifically chosen to meet the upper limit of µIVC screening methodology (section 1.3.2).

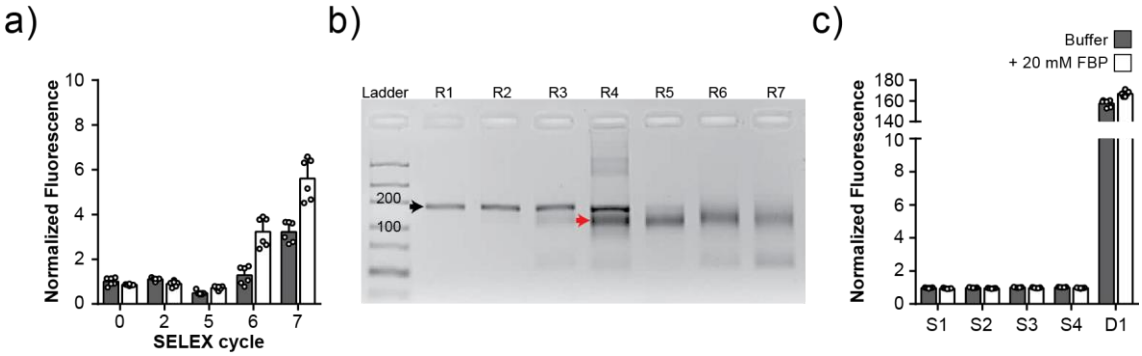


Figure 16 | Allosteric SELEX applied to the chimeric RNA library.
a, FBP-dependent fluorescence induction of the chimeric library at cycle 2, 5, 6, and 7 of allosteric SELEX compared to the unenriched library (0). Fluorescence assay performed at 1 µM purified RNA and 10 µM DFHBI in the absence (grey) or presence (white) of 20 mM FBP and normalized setting the background fluorescence of cycle 0 in the absence of FBP to 1. **b,** Ethidium bromide-stained 4% agarose gel depicting the PCR products of the seven allosteric SELEX rounds performed on the chimeric library, showing the appearance of a truncated band (red arrow) alongside the library pool PCR band (black arrow) from the third selection round onwards. **c,** FBP-dependent fluorescence induction of the top-five enriched NGS sequences in **Table 3** compared to aptamer D1. Fluorescence assay performed at 1 µM RNA, 10 µM DFHBI in the absence (white) or presence (grey) of 20 mM FBP in selection buffer (50 mM KCl, 25 mM $MgCl_2$ in 40 mM Tris pH 7.9).

3.4.2 Selection of FBP-responsive fluorogenic sequences of the chimeric library via allosteric SELEX

The chimeric library was first subjected to the previously developed allosteric SELEX protocol (Section 3.3) to identify FBP-dependent DFHBI binders.

The previously failed selection for FBP (**Figure 11 b**, **Figure 14 c**) was selected Spinach selection buffer (125 mM KCl and 5 mM MgCl2 in 40 mM HEPES pH 7.5). Here, the buffer system was changed to 50 mM KCl, 25 mM $MgCl_2$ in 40 mM Tris pH 7.9 to have the same buffer composition of the µIVC screening platform, which relies on *in vitro* transcription by T7 enzyme within one droplet, and it was kept constant not to influence this critical step of the screening (section 3.4.3).

The allosteric SELEX protocol was carried out for 7 cycles, and the enrichment of FBP-responsive fluorogenic RNAs was tested by comparing the light-up

	Enriched sequences aligned to the initial pool	RNA length	Frequency (%) R0	R5	R6	R7
Pool	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGA-GTAGAGTGTGAGCTNNNNNCACANGTGTTGACGACTCGCTGNNNNN	118				
S1	CA--------------------CTACTGTTGA-GTATGGCTCTCTC----CACACGTGTCTCGCTGTTACA--------	46	0.00	0.21	0.74	2.18
S2	CA-AGACGGTA----------------GTTGA-GCTCGCGC---------CACA--TGTGTT-ACGACT-GCTGATTT-	47	0.00	0.13	0.47	0.76
S3	--------------------GCTACTGTTGAAGTAGAGTGTGAGCTTCTA-CCCAAGTGTT--CG-CTGTAGGT----	50	0.00	0.11	0.09	0.59
S4	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGA-GTAGAGTGTGAGGTTG------------------------------	48	0.00	0.16	0.66	0.47
S5	-AGG-CGG-------AGTTCGT---ATTTGA---TAAGTGTGAGCT-------AACGTCATGTGCACTCG-TGTCTTA	55	0.00	0.07	0.34	0.42

Table 3 | Sequence trimming is revealed in NGS sequence analysis on allosteric SELEX. Most enriched sequences in NGS analysis aligned to the library pool. The conserved motifs on the sequences are highlighted (red). The RNA length and enrichment frequency per cycle are reported.

fluorescence with/without FBP throughout the selection, similarly to previous enrichment tests (section 3.2). Encouragingly, an FBP-dependent fluorescence was obtained from cycle 5 onwards (**Figure 16 a**). However, the PCR profile of the SELEX cycles displayed the co-enrichment of a by-product band – which *in vitro* selection procedures can suffer from[164] – appearing in cycle 3 and being progressively over-represented until outcompeting the library band amplification from cycle 5 onwards (**Figure 16 b**).

The enrichment of shortened sequences was also revealed by next-generation sequencing (NGS) analysis of the selection (**Table 3**). The top-5 enriched sequences identified by NGS in the last selection cycle were missing critical regions of either Baby Spinach and or C45 and exhibited an overall sequence trimming from 118 nt – length of the initial RNA library – to ~50 nt. This sequence information loss resulted in randomization of the sequence information of the chimeric library throughout SELEX progression, shown in the NGS nucleotide distribution analysis (**Supporting Figure 3 b**).

As a result, the top-5 enriched sequences identified by NGS exhibited no fluorogenic properties – with or without FBP – due to the severe sequence information loss (**Figure 16 c**). Although the RNA pool was purified before each selection cycle, the occurrence of the truncated by-products could not be prevented, and no additional selection cycles could be performed.

3.4.3 Selection of FBP-responsive fluorogenic sequences of the chimeric library via µIVC screening

Parallel to the allosteric SELEX protocol, the chimeric library was screened through a microfluidic-assisted *In vitro* Compartmentalization (µIVC) pipeline to sort single RNA species based on their fluorescence capacity under controlled conditions[114,163]. The µIVC screening was previously used to identify iSpinach, an optimized Spinach variant (section 1.3.2).

This screening platform consists of three main on-chip steps (**Figure 17 a**) that ensure the creation of a highly monodispersed – i.e., controlled drop-size[165] – RNA-in-droplet emulsion, each encompassing one single gene of the library to be screened[141]. First, the DNA library was diluted into a PCR mixture, adjusting its dilution to obtain a distinct gene/droplet ratio (λ), based on the Poisson distribution[166]. Typically, $\lambda<0.3$ is used to screen libraries of up to 10^6 variants[130,141,163] but, since the chimeric library encompasses ~ $4*10^6$ variants, the ratio was increased in the first screening cycle ($\lambda=1$, Section 5.4.3) to account for the high variability of the library. Once the correct dilution was established, water-in-oil droplets were produced by infusing the PCR mixture into a microfluidic chip with fluorinated oil phase[141] (ddPCR, **Figure 17 a**). Droplets were thermocycled off-chip to amplify each DNA variant inside the droplets and then re-injected into a droplet fusion device (Droplet fusion, **Figure 17 a**) that synchronized them one-to-one with a droplet containing *in vitro* transcription (IVT) mix supplemented FBP in case of positive screening. Pairs of droplets were fused when passing between a pair of electrodes energized by an AC field. Finally, the single RNA species in droplets were transcribed (ddIVT, **Figure 17 a**) and sorted in a fluorescence-activated droplet sorting device[116] (FADS, **Figure 17 a**) according to their fluorogenic properties. As a result, the fluorogenic property of the whole population could be quantified (**Figure 17 b**). Within the single-RNA profile, dextran-Texas Red (Orange fluorescence) was used to identify single-fused drops, while DFHBI-1T (Green fluorescence) identified sequences according to their fluorogenic output.

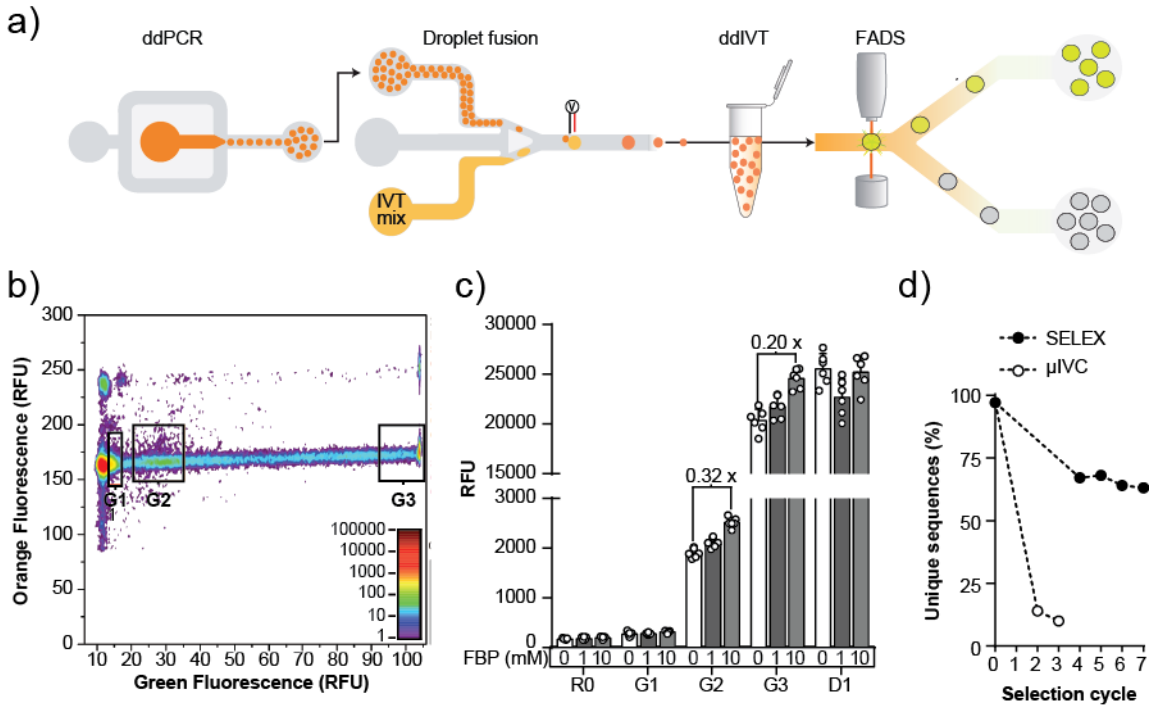


Figure 17 | An FBP-responsive pool was selected using µIVC screening.
a, Experimental workflow of µIVC to generate monodispersed RNA-in-droplet emulsion for fluorescence screening, starting from the individualization of the genes contained in a library followed by PCR amplification (ddPCR). Each PCR droplet is then fused with a larger one containing an *In vitro* Transcription mixture (IVT mix). Finally, droplets were sorted based on their fluorescence content using a Fluorescence-Activated Droplet Sorter (FADS). **b,** Fluorescence profile of the droplets obtained during the third screening round. Droplets resulting from 1:1 fusion between PCR and IVT droplets were identified thanks to the orange fluorescence (150-175 RFU) of Texas-Red added to each mixture. Concomitantly, the droplets were gated by DFHBI-1T green fluorescence according to their fluorescence readout: G1 (15-17 RFU), G2 (22-35 RFU), and G3 (93-113 RFU). For screening cycles 1 and 2 refer to **Supporting Figure 4**. **c,** FBP-dependent fluorescence induction of the RNA libraries (1 µM) obtained from the gates G1-G3 in (**b**) compared to the unenriched library (R0) and D1 aptamer (D1) at increasing concentration of FBP in selection buffer. The FBP-dependent fold-change is indicated as a numeric value. **d,** Next-generation sequence analysis revealed a substantial reduction of the number of unique sequences found in the RNA populations obtained from selection cycles 0, 2, and G2 of cycle 3 (3) compared to the allosteric SELEX performed on the same library (**Fig. 12 a**).

With this screening rationale, two initial sorting cycles were performed in the presence of FBP to enrich the library in sequences prone to fluoresce (**Supporting**

Figure 4). A third screening round was then performed in the absence of FBP to gate the droplets in nil (G1), dim (G2), and highly fluorescent (G3, **Figure 17 b**). In this way, the non-fluorogenic RNAs (G1) and those displaying FBP-independent fluorescence (G3) were separated from FBP-dependent fluorescing species (G2, **Figure 17 b**).

The three gated populations were tested by fluorescence assay measuring the light-up change in the presence of FBP (**Figure 17 c**). Expectedly, the fluorescence background increased according to the gated population, shifting from a low fluorogenic signal for G1 to high fluorescence for gate G3. More importantly, the G2-gated RNA population revealed the highest FBP-dependent increase of fluorescence (0.32-fold) compared to G1- and G3-gated populations (**Figure 17 c**) and was thus further characterized.

3.4.4 µIVC-SELEX comparison by NGS analysis

The selection performance of µIVC was investigated via NGS and compared to allosteric SELEX results obtained on the chimeric library, which also showed an increase of FBP-dependent fluorescence but with no enriched functional sequences (Section 3.4.2).

First, the capability of the two methods to deplete non-target sequences was compared (**Figure 17 d**). µIVC efficiently de-enriched non-specific sequences, with a percentile drop of 14% already from cycle 2, and to 10% in cycle 3 gate 2 (µIVC, **Figure 17 d**). Instead, the unique sequences count remained constant at 63% even after 7 SELEX cycles (SELEX, **Figure 17 d**).

Secondly, the capability of µIVC to maintain the structural complexity of the library was investigated (**Supporting Figure 5**). The chimeric library drifted to an utterly random pool with SELEX due to the enrichment of biased and trimmed sequences (**Supporting Figure 3**). Consequently, the NGS nucleotide distribution analysis showed a net loss of the structural information of Baby Spinach (nucleotides 1-46 nt) and C45 (nucleotides 52-73 nt, **Supporting Figure 3 b**). Conversely, µIVC

maintained the structural identity of the initial pool, predominantly in the region of Baby Spinach, while C45 was preserved to a lesser extent (**Supporting Figure 5 b**).

The persistence of full-length sequences, together with a reduced degeneration of the constant sequences compared to the allosteric SELEX protocol (**Table 3**), pointed at the contribution of µIVC screening in protecting the system from unwanted events, e. g., recombination during PCR and parasite dissemination[164]. Thus, the comparative NGS data seemed to indicate that µIVC pipelines based on fluorescence-based screening out-performed the affinity matrix-based SELEX approach in selecting and identifying FBP-responsive fluorescent RNA sequences.

3.4.5 Identification of an FBP-responsive sequence

To identify sensor candidates within the FBP-responsive population of cycle 3 G2, both NGS and Sanger sequencing analysis were employed (**Supporting Table 1, Supporting Table 2**).

Both analyzes revealed a similar picture of the sequence representation, with the co-enriched family of RNAs lacking a ~13 nt region (3.05-0.77%, **Supporting Table 1**) that encompassed the connecting N5 region between Baby Spinach and C45 domains, plus the 4 nt CACA stretch that was designed to hinder the folding of Baby Spinach's G-quadruplex. Also, full-length sequences were identified and showed to be enriched by NGS analysis, albeit with lower copy numbers than the shortened family (0.75-0.40% for full-length RNAs, **Supporting Table 1**).

The top-three full length RNA candidates were the sequences X (0.75%), 85 (0.74%) and 86 (0.54%). Coincidentally, 85, and 86 were also identified by Sanger sequencing (**Supporting Table 2**). These top-three NGS sequences were chosen together with 14 other candidates, picked arbitrarily amongst Sanger sequence clones, and tested for FBP-dependent fluorescence (**Figure 18**). Candidates were initially ranked based on the light-up intensity and signal reliability with FBP in

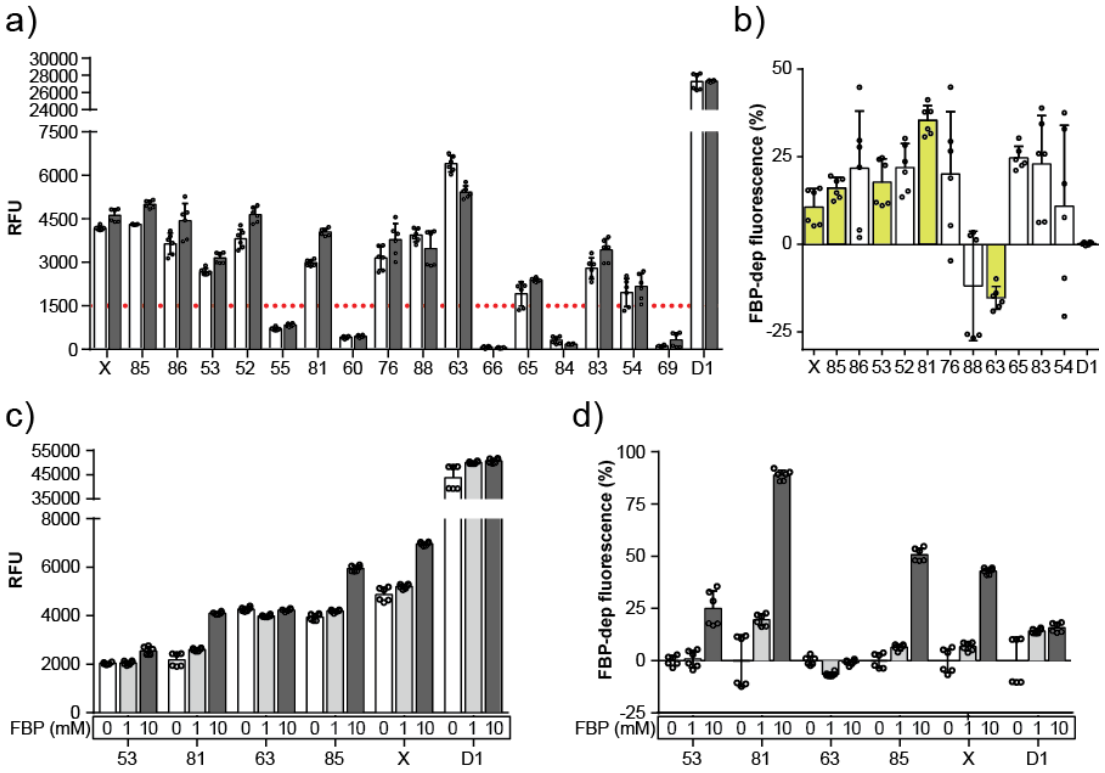


Figure 18 | S81 shows the best dose-response signal for FBP in different buffer systems.
a, FBP-dependent fluorescence induction of sensor candidates picked from Sanger and NGS analysis, compared to aptamer D1. The assay was performed in the selection buffer in the absence (white) or presence (grey) of 10 mM FBP. The threshold fluorescence was set to cut out candidates with reduced fluorescence output (red dotted line). **b,** FBP-dependent signal increase from (**a**) calculated by plotting the fold-change fluorescence increase as a percentile. Sensor candidates were picked (yellow). **c-d,** FBP-dependent fluorescence induction (**c**), and FBP-dependent signal increase (**d**) of the sensor candidates, shortlisted from **b**. The increase of fluorescence is expressed as a percentile fluorescence increase at increasing concentration of FBP. Fluorescence assay performed in minimal fluorescence buffer (40 mM Tris-HCl pH 7.9, 50 mM KCl, and 10 mM $MgCl_2$) with 1 µM purified RNA and 10 µM DFHBI.

selection buffer, which included all components of the *in vitro* transcription reaction (**Figure 18 a, b**).

Most of the clones representing the ~13 nt shortened family exhibited no fluorescence properties (66, 84, and 69, **Figure 18 a**) or were not significantly affected by the presence of FBP in the mix (83 and 54, **Figure 18 a**) and were put

aside. Instead, sequence 63 was shortlisted due to its significant decrease of fluorescence when FBP was added to the mix (-15%, **Figure 18 b**).

Together with this clone, the NGS candidates X (+10%), 85 (+16%) and the Sanger candidates 81 (+35%), 53 (+18%) were tested in minimal Tris-KCl buffer to ascertain the light-up properties of the candidates were not dependent on the buffer components, such as nucleotides or T7 RNA polymerase (**Figure 18 c, d**).

While clone 63 FBP light-down was lost probably due to the change of environment (**Figure 18 c**), sequence 81 exhibited the most robust fold-change increase of fluorescence in the presence of FBP (89 % of the maximum fluorescence at 10 mM FBP, **Figure 18 d**).

3.5 Characterization of sequence 81 as FBP light-up sensor

3.5.1 S81 responds to FBP in excess of Mg^{2+} ions

The sequence 81 (s81) exhibited the most robust light-up response to FBP among all tested candidates identified from Sanger and NGS sequencing. The ratiometric response to this metabolite was first assessed to apply this RNA as a fluorescence sensor for FBP.

The fold-change response of s81 was plotted at increasing concentration of FBP (2-600 mM) in Tris-KCl buffer containing 20, 50, and 200 mM $MgCl_2$ (**Figure 19**). Magnesium was critical for the sensor performance given that millimolar concentration of Mg^{2+} is required by both paternal aptamers Baby Spinach (5mM Mg)[94] and C45 (5-10 mM)[162]. The final concentration of Mg^{2+} in solution is, however, counterbalanced by FBP capability of chelating divalent cations *in vitro*, while it is known to coordinate Fe^{2+} in cell[167]. Consequently, Mg^{2+}-ions had to be in stoichiometric excess to FBP for s81 to fluoresce.

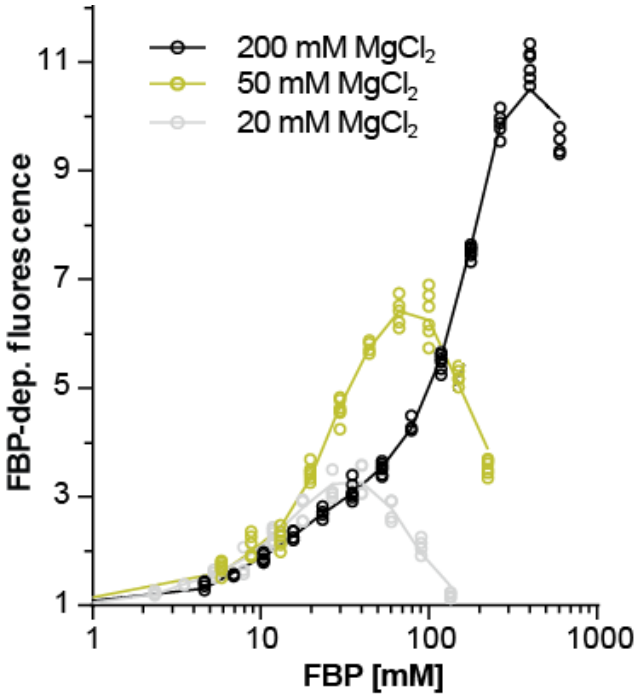


Figure 19 | S81 binds DFHBI ratiometrically to FBP and dependent on Mg^{2+} ions.

FBP-dependent induction of DFHBI fluorescence by s81 fluorescence at the indicated concentrations of $MgCl_2$. The fluorescence induction is expressed as the fold-change increment of the background fluorescence without FBP, which was set to 1. The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9, 50 mM KCl**.**

S81 exhibited an extensive ratiometric response to FBP. A ratiometric response of the sensor was detected between 2.5-25 mM FBP at 20 mM $MgCl_2$, 5-65 mM FBP at 50 mM $MgCl_2$ and 5-250 mM at 200 mM $MgCl_2$ (**Figure 19**). At these $MgCl_2$ concentrations, the linearity of detection plateaux at 40 mM, 100 mM, and 400 mM FBP, namely when the 2:1 FBP: Mg^{2+} stoichiometry is reached, and all the magnesium ions are chelated.

To note, FBP >400 mM could not be reached due to the visible precipitation of FBP: Mg^{2+} complexes. Thus, a maximal detectable concentration of FBP by s81 could not be obtained, and the EC_{50} of 81 for FBP could not be determined.

Nevertheless, s81 demonstrated a ratiometric fluorescence response to FBP in the millimolar range insofar that magnesium ions were in excess, making 81 a sensor for FBP easily applicable *in vitro* and potentially in cells.

3.5.2 Sequence 81 sensing mechanism relies upon G-quadruplex destabilization

The ratiometric response to the millimolar concentration of FBP makes s81 a sensor for this metabolite, but its structure-switching mechanism had to be investigated and could be different from the allosteric design of the chimeric library, from which 81 was identified. The chimeric library design adopts the sequence-sequestering strategy to promote allosteric regulation and targets critical nucleotides in the G-quadruplex domain (Section 3.4.1).

To obtain evidence that s81 de-structures the G-quadruplex domain, the fluorescence response to FBP was tested at increasing concentration of K^+ ions; these ions are known to stabilize G-quadruplex domains (**Figure 20**). It is well-described in the literature that Spinach structure requires 2 K^+-ions for the folding of the two G-quartets. Warner *et al.* have shown the coordination map of these ions embedded in the G-quartet of Spinach's crystal structure[94]. Additionally, Ageely *et al.* detailed that – at ≥50 mM of alkali metal salts – Baby Spinach exhibited strong fluorescence only with potassium and not rubidium, sodium, cesium, or lithium chloride[147].

The increase of K^+-ions negatively impacts 81's performance by increasing its background fluorescence (black bars, **Figure 20**) while curbing the light-up detection signal for FBP (yellow bars, **Figure 20**). Specifically, the background fluorescence exhibits a 10-fold increase at 400 mM KCl compared to 50 mM KCl, in which the s81 was selected (**Figure 17**). When 40 mM FBP are added, s81's fluorescence robustly increases at 50 mM KCl (4.76-fold, **Figure 20**) but progressively curbs at a higher concentration of K^+-ions, reaching its lowest point at 400 mM KCl (1.31-fold, **Figure 20**).

Likewise, the increase of K^+-ions benefits the fluorescence of D1 aptamer control, but to a minor extent (**Supporting Figure 6**). Specifically, D1's fluorescence is 0.25-fold in 400 mM KCl compared to 50 mM. Notably, the addition of FBP is

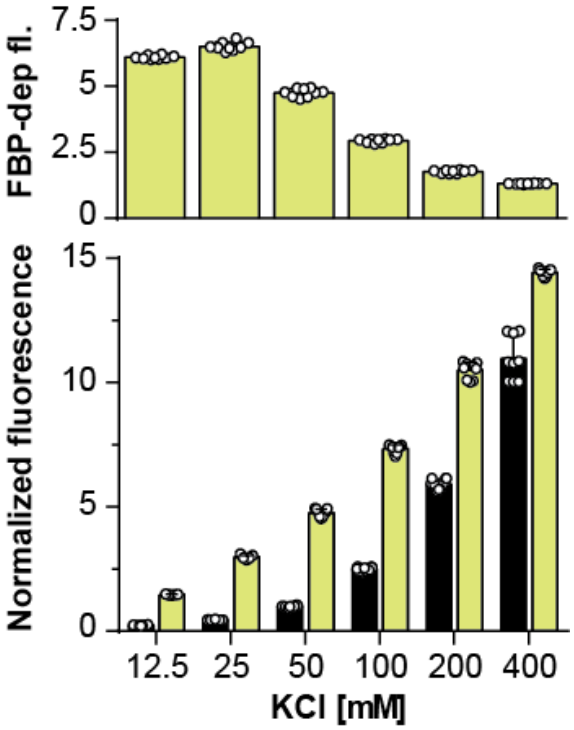


Figure 20 | K^+ ions stabilize Spinach's G-quadruplex and decrease s81 sensing performance. The bottom graph plots the fluorescence readout measured in the absence (black) and presence of 40 mM FBP (yellow) at increasing concentration of KCl. Values were normalized by setting to 1 the background fluorescence at 50 mM KCl. The top graph plots the fold-change increment of each condition. The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9 and 30 mM $MgCl_2$.

always detrimental for D1's fluorescence (black vs. yellow bars, **Supporting Figure 6**).

Thus, the FBP-dependent light-up response obtained with s81 and not D1 control confirms the specificity of the interaction. Furthermore, the loss in the light-up performance of s81 for FBP at increasing KCl (from 4.76- to 1.31-fold) is in accordance with the model of G-quadruplex destabilization caused by conjoining C45 aptamer in an allosteric manner.

By increasingly stabilizing the G-quadruplex, K^+ ions counterbalance the allosteric control that C45 has on Baby Spinach, causing the background fluorescence of the sensor to increase (bottom, **Figure 20**) and its FBP-dependent fluorescence to decrease (top, **Figure 20**). However, the decrease of fluorescence reported by D1 control by FBP could indicate a parallel unspecific destabilization caused by the metabolite (**Supporting Figure 6**), but this effect is negligible at 50 mM KCl.

3.5.3 Mutational analysis on the CACA-stretch

The K^+-shift assay (**Figure 20**) indicated that s81 sensing mechanism adopts the allosteric regulation of its G-quadruplex folding, as initially designed in the chimeric library. To confirm that the library design is maintained in the s81, series of mutants of the CACA region were designed to weaken its hybridization to the residues 62-65 nt (UGUG, **Figure 21**) of Baby Spinach's G-quartet.

As expected, little as a single point mutation in the CACA region almost halved (from 4.5 to 2.5-fold for 1PM, **Figure 21**) the performance of the sensor and additional mutations (2PM, 4PM) lead to a complete loss of allosteric control. Of note, the background fluorescence linearly increases the more nucleotides in the

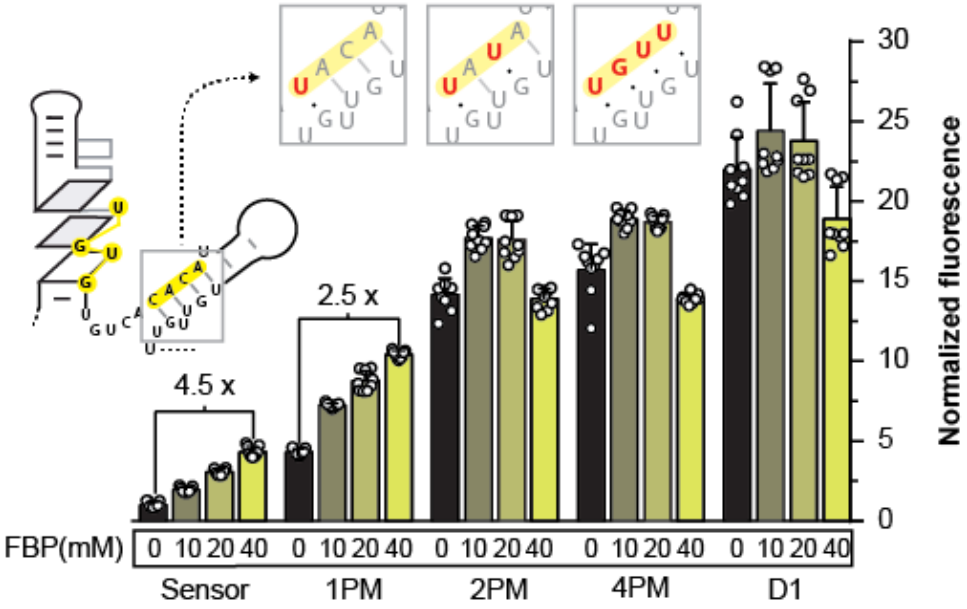


Figure 21 | Point mutant analysis in the CACA region of s81.
Dose-response analysis of the FBP-dependent fluorescence of s81 fluorescence compared to the point mutants of the CACA region. The modified nucleotides in each mutant are shown. The fold-change response for 40 mM FBP from s81 (Sensor) and the single-point mutant (1PM) are indicated. The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9, 50 mM KCl, 30 mM $MgCl_2$. Values were normalized, setting to 1 the background fluorescence of the sensor without FBP.

CACA region of s81 are mutated (black bars, **Figure 21**) and reaches the fluorescence level of aptamer D1 when four nucleotides are mutated (4PM, **Figure 21**).

The decrease of the allosteric control concomitantly with the increase of the background fluorescence obtained by mutating the CACA-stretch indicated that, according to the initial design strategy, this region of s81 is critical to transduce FBP binding into a fluorescence output through Baby Spinach's G-quartet destabilization.

3.5.4 EC_{50} measurement for DFHBI

Based on the previous dose-response assays, the optimal ion concentration for s81 was set to 50 mM KCl (**Figure 20**) and 200 mM $MgCl_2$ (**Figure 19**) to have the maximal detection limit to FBP. Then, this ion setting was used to investigate the concentration-dependent binding of s81 to DFHBI in the presence and absence of FBP (**Figure 22**).

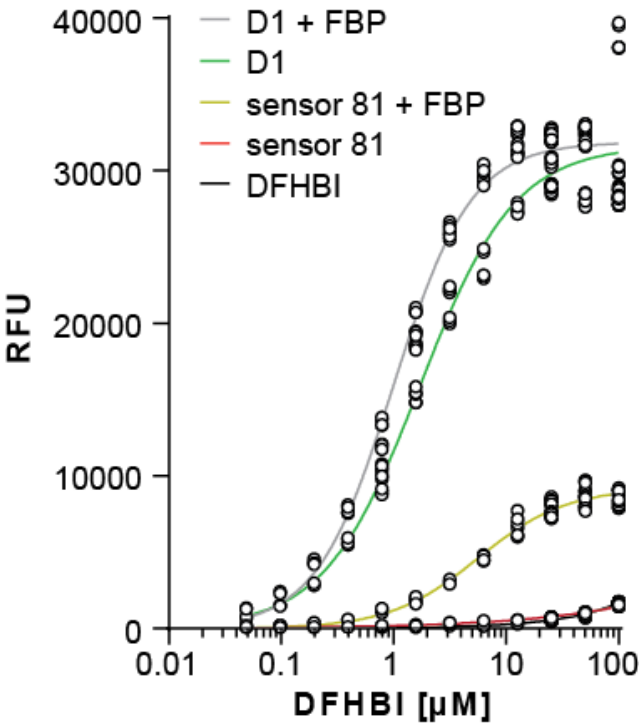


Figure 22 | S81 affinity to DFHBI depends on FBP.
EC_{50} measurement of s81 towards DFHBI in absence or presence of 250 mM FBP, compared to D1 aptamer and DFHBI autofluorescence in 40 mM Tris-HCl pH 7.9, 50 mM KCl, 200 mM $MgCl_2$. The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9 and 30 mM $MgCl_2$.

In this buffer, the s81 background fluorescence overlapped with DFHBI autofluorescence. When 250 mM FBP were added, namely the highest point in the linearity range of detection obtained in **Figure 19**, a concentration-dependent interaction with DFHBI was obtained with an EC_{50} value of 5.56 ± 0.40 µM. Instead, D1 control showed comparable interaction with DFHBI in the absence (EC_{50}: 1.61 ± 0.10 µM) or presence (1.13 ± 0.07 µM) of FBP. By comparing the interaction curve of s81 to D1 aptamer control, the EC_{50} for 81 was 4.9-fold lower with FBP, whereas the end-point fluorescence signal was 3.3-fold less intense.

Ultimately, these results confirmed that s81 binding to DFHBI is allosterically-dependent on FBP, but its affinity to DFHBI is inferior to the parental aptamer Spinach.

3.5.5 Specificity assay for glycolytic intermediates

Last, the specificity of s81 was investigated by measuring the dose-response fluorescence increase of fructose 1,6-bisphosphate compared to a panel of molecules structurally related to this metabolite (FBP, **Figure 23 b**).

Being the metabolic precursor of FBP in glycolysis, fructose 6-phosphate (FMP, **Figure 23 c**) and glucose 6-phosphate (G6P, **Figure 23 d**) were used to test the capability of s81 to distinguish chemically-similar structures and understand if this sensor would signal only FBP inside cells. Instead, the loosely related glucosamine 6-phosphate (GluN6P, **Figure 23 e**) and glucose (Glu, **Figure 23 f**) were analyzed to establish whether the sensor can specifically target glycolysis and glycolytic metabolites.

FBP was found to elicit the highest fluorescence response of s81 at the tested concentrations (1.2-, 1.8- and 5.3-fold at increasing FBP), indicating the specificity of recognition for this key-metabolite (yellow bars, **Figure 23 e**). FBP signal was twice more specific than for closely related metabolites, i.e., FMP and Glu6P (1.1-, 1.4- and 2.6-fold). Encouragingly, GluN6P and Glu showed little to no light-up

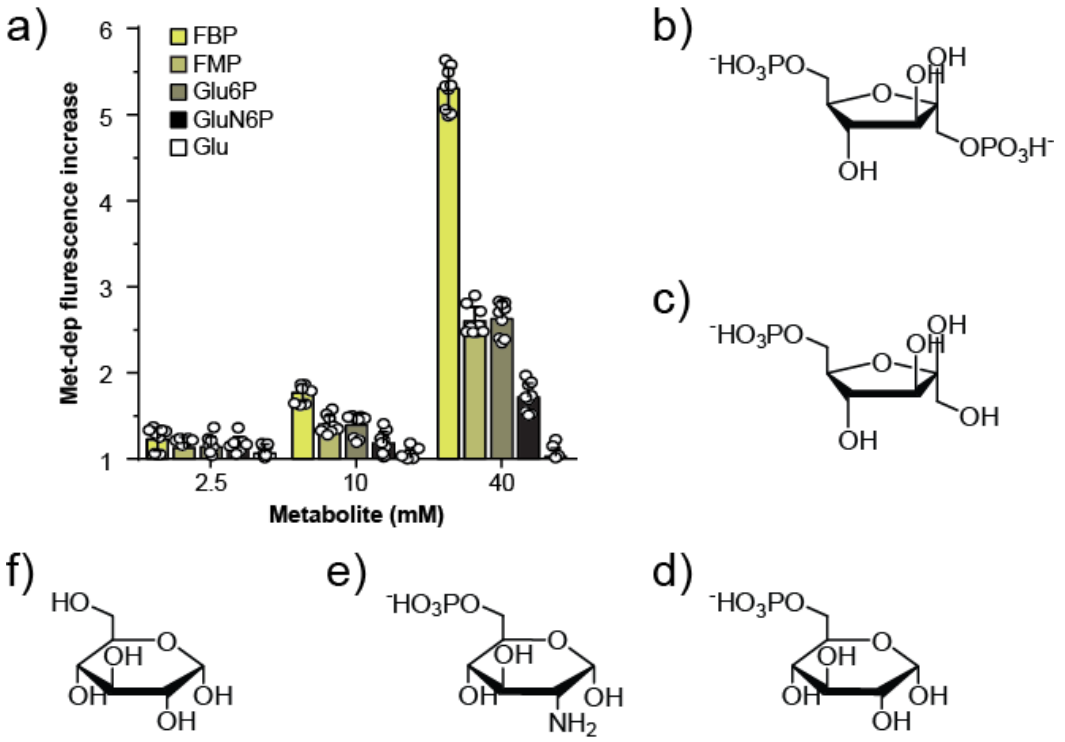


Figure 23 | S81 preferentially responds to FBP compared with other structurally similar metabolites.
a, The fluorescence emission of s81 measured at the indicated concentrations of Fructose 1,6-bisphosphate (FBP, **b**), Fructose 6-phosphate (FMP, **c**), Glucose 6-phosphate (Glu6P, **d**), Glucosamine 6-phosphate (GluN6P, **e**) and Glucose (Glu, **f**). Data normalized to the background fluorescence of s81 and to the unspecific fluorescence variation of aptamer D1 with the different metabolites, which were set both to 1. The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9, 50 mM KCl, 30 mM $MgCl_2$.

signal, signifying that the sensor only partially recognized the metabolic precursors of FBP, namely G6P and FMP, because of their molecular similarity to our target analyte.

3.5.6 NGS analysis repetition of the µIVC screening on the chimeric library and motifs identification

S81 showed the best FBP-response in the comparative analysis of a panel of sensor candidates (**Figure 18**). Despite its performance, s81 was only identified by

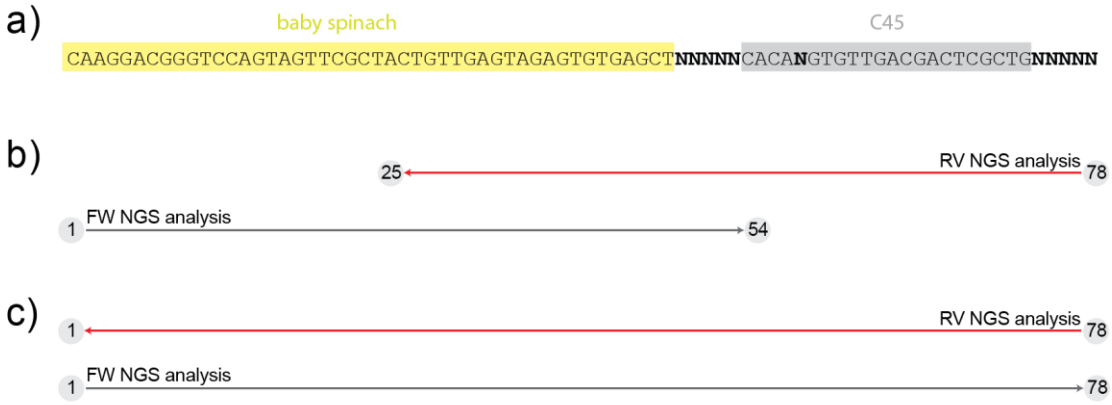


Figure 24 | Sequence coverage of the chimeric library between the 75 and the 150 NGS sequencing kit.
Schematic representation of the sequence coverage of the chimeric library (**a**) when the 75 cycle (**b**) or the 150 cycle (**c**) NGS sequencing kit was applied. The baby spinach (yellow) and C45 (grey) domains within the chimeric library are highlighted. **b,** When the 75 cycle kit was used, the library could not be sequenced in its entirety from both the FW- and RV-sides. Nevertheless, the sequence overlap between FW and RV (25-54 nt) was used to manually identify the most enriched sequences, shown in **Supporting Table 1**. **c,** The whole library could be sequenced from both FW- and RV-sides when the 150 cycle kit was used.

SANGER sequencing, not by NGS analysis on the enriched pool 3.G2 (**Supporting Table 1, Supporting Table 2**)**.** However, it is not uncommon that sequences with an aptamer function are not to be found in the most enriched sequence pool[150,168], likely due to propagation bias of the PCR during the selection algorithm[164,169].

Nevertheless, the inability to identify s81 raised the question of whether the sequencing-depth in the performed NGS analysis was sufficient, which was caused by two issues: the sequencing kit used (**Figure 24**) and the enrichment of the truncated family of sequences throughout the screening (**Supporting Table 1, Supporting Table 2**). First, the 75 cycles NGS sequencing kit did not permit to cover the whole library from both the forward (FW) and reverse (RV) primer sites, as shown in the schematic NGS coverage scheme in **Figure 24 b**. Nevertheless, the sequence information was obtained due to a partial overlap of the FW- and RV-side analysis (25-54 nt overlap, **Figure 24 b**). Thus, the most enriched hits from

both NGS sides were manually overlapped, and the most enriched sequences could then be identified (**Supporting Table 1**). However, only the most enriched sequences could be identified with this method, and no deep-sequencing investigation could be performed. Furthermore, the enrichment of the ~13 nt shortened family further confounded the analysis; this family appeared to have no allosteric properties (**Figure 18**). Its appearance concealed the identification of full-length sequences shown in the NGS enrichment table for cycle 3.G2 in **Supporting Table 1**.

Given the previous analysis limitation, the NGS analysis was repeated, and the previous issues were circumvented. First, the 150 cycle NGS sequencing kit was used to cover the whole library length from both forward and reverse sequencing sides (**Figure 24 c**). The obtained sequences were mined for the consensus sequence of the chimeric library to bypass the enrichment of the shortened sequence family (AACAAGGACGGGTCCAGTAGT-TCGCTACTGTTGAGTAGA-GTGTGAGCTNNNNNCACANGTGTTGACGACTCGCTGNNNNN) and isolate the full-length sequences that were enriched in the µIVC screening. To note, every screening cycle was sequenced in the second NGS analysis to have a clear understanding of the evolution of s81. **Supporting Table 3** lists the top-fifty sequences of cycle 3.G2 with enrichment >0.1 %, and their percentile frequency in every screening cycle are listed in **Supporting Table 4**.

From the comparison of the sequencing data with the previous analysis, it could be noted that three of the top-five enriched sequences were previously identified. The most enriched sequence A1 (12.8%, **Supporting Table 4**), which was not previously identified, was followed by sequences 53, X, and 85 (11%, 8.9%, and 8%, respectively **Supporting Table 4**). Also, sequence 86 was re-identified as the top-8th enriched sequence (4.5%, **Supporting Table 4**). The partial overlap between this and the previous sequencing indicated that, albeit the previous lacked sequencing depth, the most enriched sequences could be identified regardless.

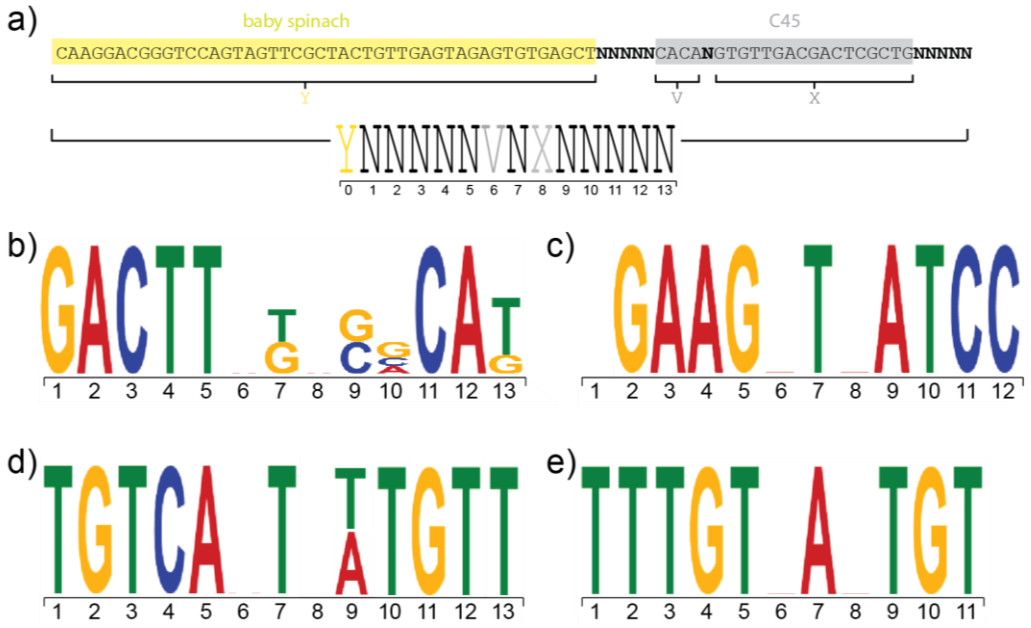


Figure 25 | Motifs identified from MEME analysis on 3.G2 pool top-50 enriched sequences.
a, The constant regions of the chimeric library were tagged to isolate the random regions of the library. The chimeric library's baby spinach region (yellow) was tagged as Y in position 0, whereas the two constant regions of C45 (grey) were tagged as V in position 6 and X in position 8. So designed, the chimeric library counts 13 positions, of which 11 were randomized. **b-e,** Conserved motifs identified by MEME analysis on the sequences listed in **Supporting Table 3**. Motif 1 (**b**) was shared between sequences A2 (also 53), A31, A19, and A7. Motif 2 (**c**) was shared between sequences A3 (also X) and A45. Motif 3 (**d**) was shared between sequences A48 (also 81) and A21. Motif 4 (**e**) was shared between sequences A30, A20, and A15. The motifs are color-coded according to the nucleotide, and the positions are consistent with the library consensus shown in **a**.

Contrary to the previous analysis, however, s81 could be identified in the in-depth sequencing as the top-48th enriched sequence in cycle 3.G2 (0.13%, **Supporting Table 4**). The low enrichment of s81 was expected since it was not promptly identified in the previous NGS analysis. Nevertheless, its enrichment profile throughout the µIVC screening was surprising: s81 was strongly enriched during the positive selection screens in the presence of FBP (0.01%, and 0.78% in cycle 1-2 respectively, **Supporting Table 4**), but it was de-enriched in all gated pools of

screen 3 that were performed in the absence of FBP (0.01%, 0.13, and 0.27% in gates 1-3 respectively, **Supporting Table 4**).

In comparison, the enrichment of 53, X, 85, and 86 – that were also tested for their FBP-response but exhibited a lower sensing performance than 81 (**Figure 18**) – exhibited a similar profile up to the second screening cycle. However, they were instead enriched explicitly in the 3.G2 pool, which showed the best FBP-response in the fluorescence test of the various pools (**Figure 17 c**). From these results, it could be reasoned that the bulk of the FBP-response obtained from the 3.G2 pool was due to the enrichment of sequences like 53, X, and 85.

Further, the sequencing results were analyzed with the MEME suite[170] in search of reoccurring motifs between the sequences listed in **Supporting Table 3**. The constant regions of the library were tagged (**Figure 25 a**) to ensure that the motif would be identified only from the randomized part of the chimeric library. Encouragingly, s81 was identified as a motif (motif 3, **Figure 25 d**) together with

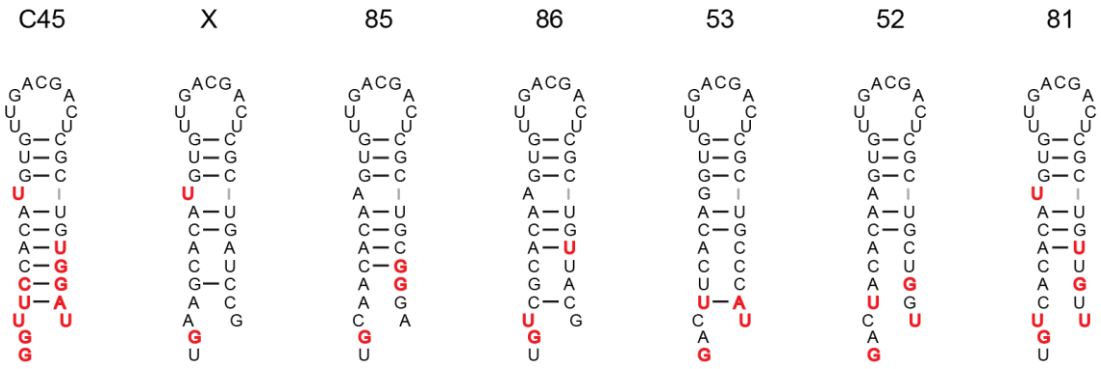


Figure 26 | 81 exhibited the highest sequence similarity to C45 amongst all sensor candidates
The truncated version of the C45 aptamer, which was embedded in the chimeric library, was overlapped with the sequence of the sensor candidates shortlisted in **Figure 18**. Sequences are ordered from the least (X) to the most similar (81) to C45. The nucleotides of C45 randomized in the chimeric library are highlighted in red, together with the similar nucleotides to the sensor candidates.

other two motifs similar to 53 (motif 1, **Figure 25 b**) and X (motif 2, **Figure 25 c**). Also, a T-rich motif was enriched (**Figure 25 e**), but no similarity was found to the previously tested sensor candidates.

The finding that s81 was conserved in motif 3, beside one nucleotide (position 9, **Figure 25 d**), might indicate that the sequence identity is an essential sensor feature. A comparative analysis of the previously tested sensor candidates (**Figure 18**) identified 81 as the sequence with the most similarity to the paternal C45 aptamer (6 nt overlaps of the random region, **Figure 26**). Therefore, while 81 exhibited low enrichment in the selection, its superior performance in the FBP-response assay to other sensor candidates could be explained by a balance between newly selected nucleotides and, on the other hand, a remarkable similarity to the paternal C45 aptamer.

3.6 Application of s81 to image oligomycin-induced glycolytic response in HEK293T

3.6.1 Expression of s81 as a circular scaffolded RNA

The *in vitro* characterization of 81 highlighted its ratiometric response to FBP (**Figure 19**), specificity (**Figure 23**), and elucidated the structure-switching mechanism of the sensor (**Figure 20- Figure 22**). Next, the sensor was expressed as a fusion construct to an RNA scaffold to test the sensing performance of s81 in the cellular milieu of mammalian cell lines.

To embed the RNA of interest into the structure of a protective RNA scaffold is a conceptually simple and established strategy to create a stably-expressed functional RNA protected by cleavage in the cell[115,133]. Initially, the $tRNA^{Lys}_3$ scaffold published by Ponchon *et al.*[171] allowed to enhance the brightness and stability of fluorogenic RNAs, especially in bacteria[77,121,133] and to a lesser extent for mammalian cells[97]. Later, Filonov *et al.* demonstrated that the use of this tRNA induced RNA cleavage in mammalian cells - counteracting the very purpose of the

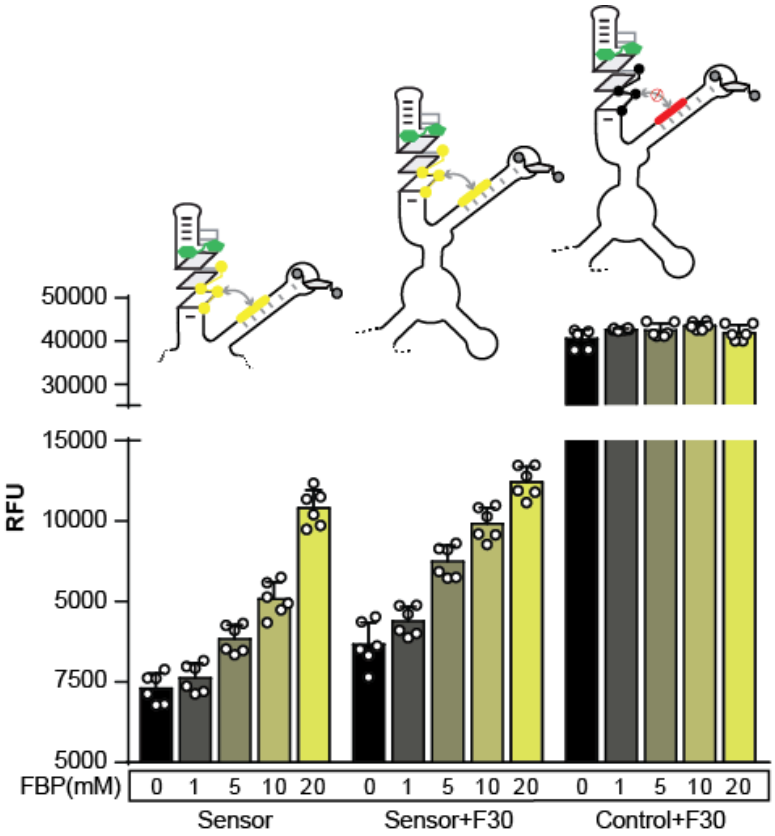


Figure 27 | FBP light-up sensor is not affected by the integration in the F30 scaffold. FBP-dependent fluorescence induction of increasing concentration of FBP of s81 compared with the F30 integrated version of the sensor (Sensor+F30) and 4PM control (Control+F30), showing comparable light-up response towards FBP for both versions of s81. Conversely, 4PM + F30 control keeps showing no allosteric light-up response. Fluorescence assay performed in 40 mM Tris-HCl pH 7.9, 50 mM KCl, 50 mM NaCl and 20 mM $MgCl_2$. Fluorescence assays were performed with 1 µM purified RNA and 10 µM DFHBI.

scaffold – and identified a better suiting scaffold for mammalian expression systems, a re-engineered version of the Φ29 scaffold named F30[112].

Thus, F30 was used as a scaffold to enhance the stability of 81 for cellular expression. However, fusing a bulky moiety such as F30 might negatively influence the allosteric properties of the sensor towards FBP by hijacking the native folding of the RNA. Therefore, the light-up FBP induction of s81 embedded into the F30 scaffold was tested *in vitro* (**Figure 27**). The fluorescence assay confirmed that the

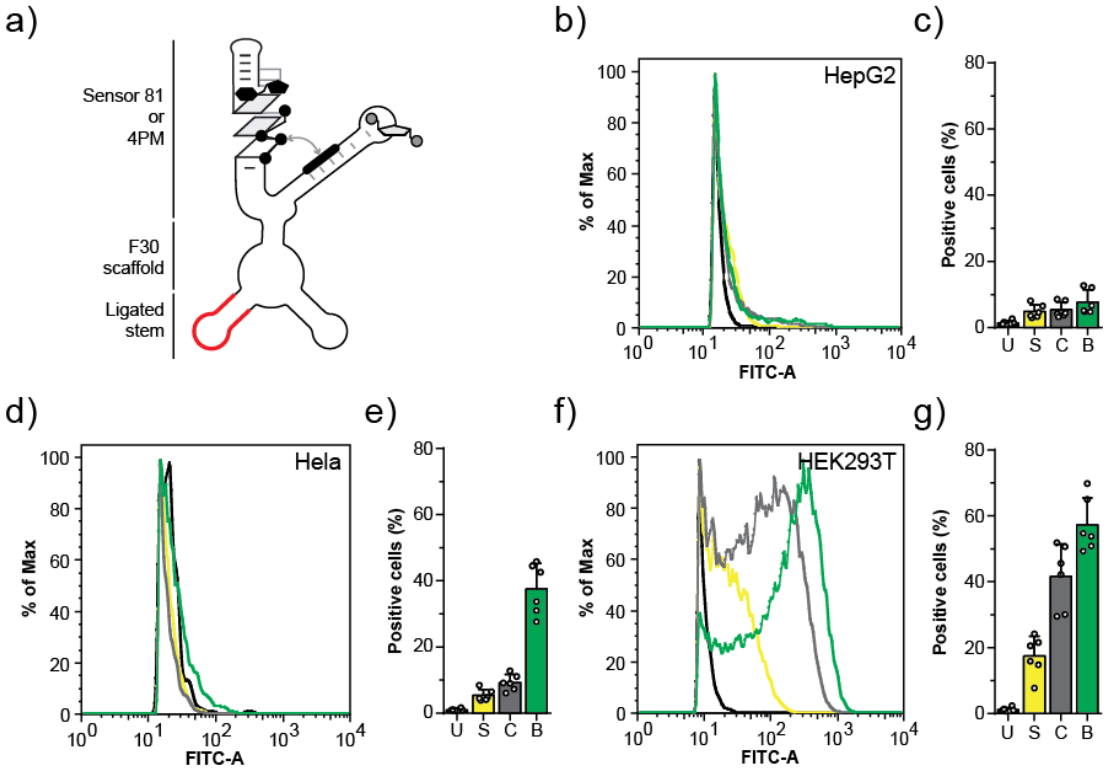


Figure 28 | Expression of s81 in mammalian cell lines as a scaffolded circular RNA. **a,** Schematic representation of the processed circular RNA that is obtained expressing s81 and 4PM control intracellularly with the aid of the Tornado system[135]. The circularizing stem at the base of the F30 three-way junction is highlighted (red). **b- g,** Cytogram plot and positive cells percentile of HepG2 (**b,c**), Hela (**d,e**) and HEK293T (**f,g**) cell lines transiently expressing s81 (S), 4PM (C) and Broccoli (B) RNAs normalized to untransfected control (U) which was set to 1%.

light-up FBP induction is not affected by the integration, shown by the comparable fluorescence induction of s81 with or without the scaffold (Sensor vs. Sensor+F30, **Figure 27**).

The CACA-stretch 4 point mutant (4PM, **Figure 21**) was tested to ensure it would remain irresponsive to FBP after the integration and confirmed that its fluorescence was still unaffected by FBP (Control+F30, **Figure 27**). Based on these data, s81 could be expressed in the cellular system embedded into the F30 scaffold with no effect to its FBP-response and 4PM+F30 sequence – which showed no allosteric

light-up response to FBP – could be used as a valid internal control for our sensor system. Next, the FBP sensor fused to F30 was expressed as circular RNA in different mammalian cell lines using the Tornado expression vector established by Litke *et al.*[135]*.*

In the Tornado vector, the scaffolded RNA is flanked by Twister ribozymes, which undergo autocatalytic cleavage upon expression. Once cleaved, the RNA termini are ligated by ubiquitous cellular ligases[135]. The result is a highly-stable, circular RNA expressed intracellularly (**Figure 28 a**). The FBP s81 was cloned into the Tornado vector and was transiently expressed into HepG2, Hela, and Hek293T cells (**Figure 28 b-g**), and its expression efficiency was compared to 4PM and Broccoli internal controls.

The Broccoli aptamer is another DFHBI-binding RNA previously used to establish the Tornado expression system[135] (Section 1.2.2) and was used here to monitor the transfection efficiency in the various cell lines.

HepG2 cells exhibited low transfection efficiency, displaying a <10% positive cells increase for Broccoli transfection control (**Figure 28 c**), whereas Hela showed higher efficiency (~40% Broccoli signal, **Figure 28 e**) but a reduced fluorogenic signal for s81 and 4PM control (~5 and 9% respectively, **Figure 28 e**). Conversely, HEK293T cells exhibited high transfection efficiency (~56% Broccoli signal, **Figure 28 g**) and a ratiometric signal decrease from 4PM control (~42%, **Figure 28 g**) to s81 (~18%, **Figure 28 g**), which matched the fluorescence output of these RNA constructs *in vitro* (**Figure 27**).

To ensure that the sensor is functional in the cellular context, the fluorescence discrepancy between the FBP sensor and 4PM is critical. The fluorescence signal given by 4PM can be considered, in fact, as the signal that the sensor would gain if all its RNA sequences were bound to FBP and the G-quadruplex destabilization was minimal. For this reason, HEK293T was used as a model cellular system for its high-transfectability and fluorescence discrimination between Sensor and 4PM control (**Figure 28 g**).

3.6.2 Discrimination of oligomycin-induced glycolytic boost in HEK293T

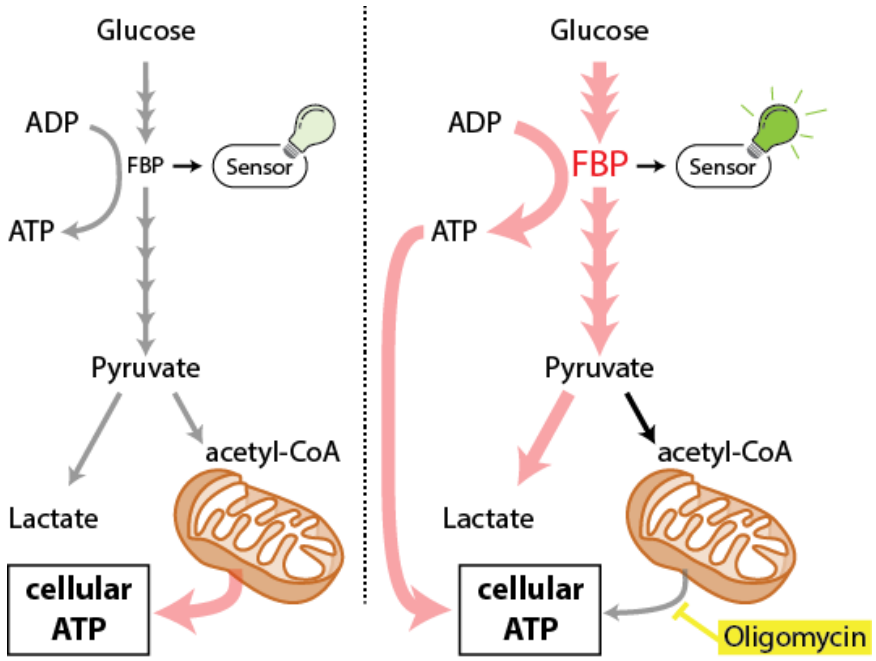


Figure 29 | Oligomycin A-induced glycolytic flux increase.
Schematic of the intracellular FBP increase in HEK293T cells caused by the oligomycin-induced glycolytic flux shift from physiological (left) to treated (right). In physiological conditions, the ATP production burden is assigned to oxidative phosphorylation, based in the mitochondria, and the ATP-production associated with glycolysis is low. The oligomycin A treatment abolishes the oxidative phosphorylation and shunts the ATP-burden to glycolysis, causing the glycolytic flux to boost and FBP concentration increase.

After identifying HEK293T as a suitable cellular model (**Figure 28 g**), in which s81 and its control 4PM exhibited comparable fluorescence properties to the *in vitro* system (**Figure 27**), the last step was to test s81's capability of image FBP in the cell context.

Yet, at large, the physiological concentration of FBP in mammals is ≤1mM[172,173], which is too low to be detected by s81. However, the intracellular FBP has been consistently shown to reflect glycolytic flux in different organisms, including specific

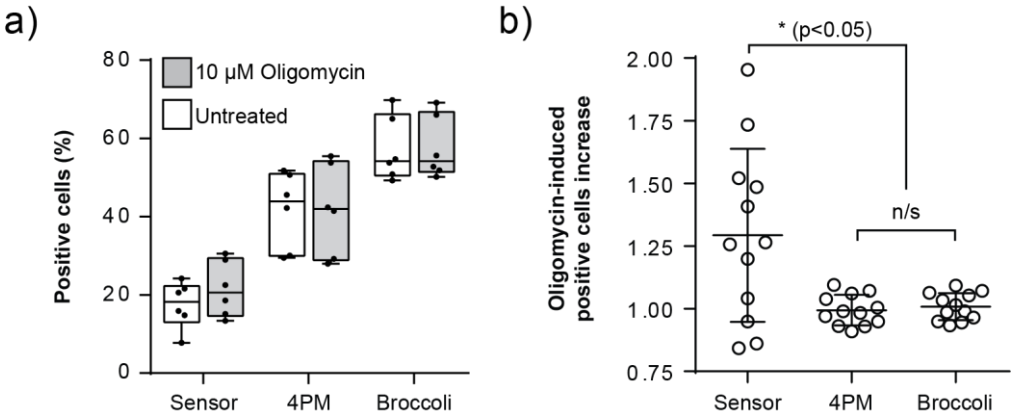


Figure 30 | S81 and not 4PM nor Broccoli responded to oligomycin A treatment in HEK293T transiently-transfected cells.
a, Positive cell percentile of HEK293T cells expressing sensor, 4PM, and Broccoli RNAs treated (grey) or untreated (white) with 10 µM oligomycin. Positive cells percentile values were obtained normalizing to untransfected cells control, which was set to 1%. Data represents n=3 independent experiments. These data were used to calculate the oligomycin-induced positive cells increase for each experimental repetition (**b**), plotting the fold-change increase for each RNA in the presence of oligomycin. The p-values were calculated using the Mann-Whitney nonparametric test.

mammal cell lines (Section 1.1.4). The FBP-glycolysis correlation could be used to test s81 using specific antibiotic molecules that boost glycolysis[174] and potentially increase the intracellular FBP concentration. Amongst glycolytic booster, the antibiotic oligomycin A has been widely used to study glycolytic flux regulation[15] and was, in particular, applied to establish the maximum rate of glycolysis that can be achieved by a cultured cell line, which is defined as glycolytic capacity[33,175,176].

Oligomycin A is a potent inhibitor of the mitochondrial ATP synthase, a crucial enzyme for oxidative phosphorylation of ADP to ATP, and its addition to cultured medium halts the physiological production of ATP in cells[174]. When oligomycin A is present, the ATP demand cannot be met by oxidative phosphorylation and is shunted to glycolysis, causing a boost in the glycolytic flux (**Figure 29**). The oligomycin treatment was combined with the developed FBP s81 to investigate whether a boost in the glycolytic rate caused a corresponding increase in the fluorescence of the sensor (**Figure 30**).

To test the capability of the sensor to discriminate the glycolytic boost, HEK293T cells were transfected with s81, 4PM and Broccoli RNAs as experimental controls to detect the fluorescence signal increase upon oligomycin treatment by flow cytometry. Changes in the fluorescence signal were identified by calculating the percentile of positive cells, namely the percentage of cells with fluorescence higher than the top 1% fluorogenic cells from the untransfected sample (**Figure 30 a**). The gating strategy used in the assay is depicted in **Supporting Figure 8**. As already highlighted in the previous transfection test (**Figure 28 g**), the fluorescence median increased gradually from s81 to 4PM control and finally to Broccoli (**Figure 30 a**). Albeit timidly, the median and interquartile range of the sensor was affected by the oligomycin treatment, whereas 4PM and Broccoli controls remained constant. Data were normalized for intra-experimental variation by calculating the oligomycin-induced positive cell increase, shown in **Figure 30 b**.

The normalized values were analyzed by Mann-Whitney nonparametric test[177], confirming significant discrimination from the untreated to the oligomycin-treated cells expressing s81 (p-value: 0.0175, **Figure 30 b**), and not 4PM and Broccoli controls (ns, **Figure 30 b**). Specifically, the non-significant change in the Broccoli and 4PM mutant fluorescence (ns, **Figure 30 b**) confirmed that s81 signal variation was not due to an increase in the RNA level, measured by Broccoli, nor to un unspecific variation in the Spinach reporting system, which would have also been shown in the 4PM control. The results were further confirmed in a double-blind setting (**Supporting Figure 9**).

Based on these data, it could be concluded that s81 specifically responded to oligomycin-treatment in HEK293T. Given the widely established correlation between oligomycin-treatment and glycolytic flux increase[33,175,176], this result indicated that the sensor had the potential for imaging the glycolytic boost caused by the antibiotic. On the other hand, the cross-specificity of s81 for other glycolytic intermediates, shown in **Figure 23**, raised uncertainties as to which metabolite –

FBP, G6P, or FMP – was recognized in the cellular context or if even a combination of these targets.

3.7 Quantification of oligomycin-induced intracellular FBP variation and assessment of s81 specificity

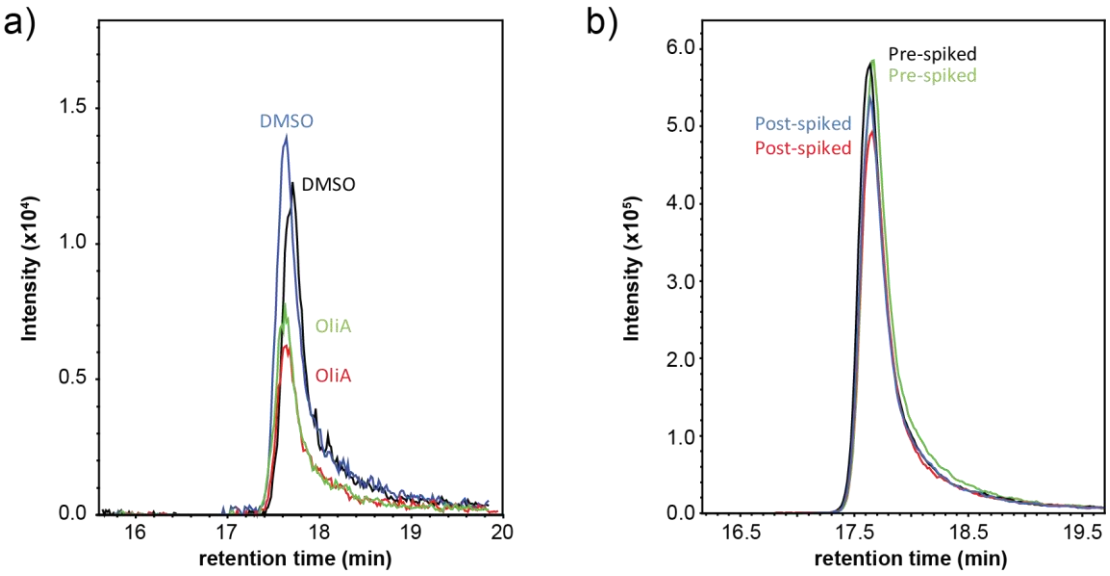


Figure 31 | Intracellular FBP decreased in HEK293T cells treated with oligomycin A. **a,** LC-MS analysis of intracellular FBP levels in HEK293T cells treated with oligomycin A (OliA) or DMSO solvent vehicle (DMSO). **b,** A pre-spiked and post-spiked control with 1 mM FBP was applied to ensure the extraction method would not significantly reduce analyte levels and to have a reference for the retention time of FBP. Each condition was measured in duplicate, and each is shown as a separated plot.

To demonstrate the potential of s81 to image the boost in intracellular FBP caused by oligomycin A, the level of this metabolite in HEK293T cells treated by oligomycin A and DMSO baseline control were quantified by LC–MS[178]. Previously, the FBP-glycolysis correlation was studied and established by Rabinowitz and co-workers in mammalian iBMK cells using oligomycin A[15,157]; cell handing, metabolite extraction, and quantification strictly followed their protocol but were applied in HEK293T cells instead.

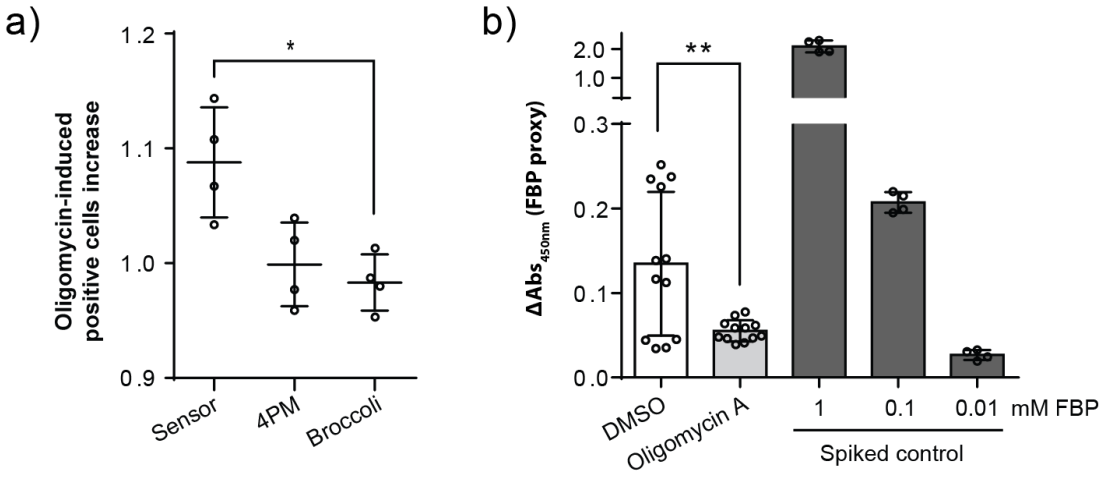


Figure 32 | Head-to-head comparison of oligomycin A to s81 signal and intracellular FBP in HEK293T cells

a-b, The flow cytometry evaluation of s81 response to oligomycin A (**a**) was repeated head-to-head to the correlated effect on the intracellular FBP level, quantified by colorimetric assay (**b**). **a,** The flow cytometry assay results were obtained as in **Figure 31**, where the positive cell percentile of HEK293T cells expressing sensor, 4PM, and Broccoli RNAs and treated with 10 µM oligomycin was normalized to the DMSO control, thus obtaining the oligomycin-induced positive cells increase. **b,** The colorimetric kit is based on the conversion of FBP to FMP by the enzyme FBPase. FMP then reacts with a chromophore forming a product with a stable signal measured at OD=450 nM. To measure only FBP converted to FMP and not the intracellular FMP, each sample's cellular baseline was measured and subtracted, obtaining the ΔAbs values plotted in the graph. The p-values were calculated using the Mann-Whitney nonparametric test.

The LC-MS picks for FBP are shown in **Figure 31**. Unexpectedly, intracellular FBP plummeted upon treatment with oligomycin A compared to the DMSO control (**Figure 31 a**). To ensure that FBP was not degraded during the extraction method, two spiked controls were tested at the start (pre-spiked) and the end of the extraction protocol (post-spiked) with 1 mM FBP, and they confirmed no metabolite was lost throughout the extraction (**Figure 31 b**). Similarly, glucose 6 phosphate and fructose 6 phosphate, the two glycolytic precursors of FBP that cross-reacted in the specificity assay for s81 (**Figure 23**), decreased upon oligomycin A treatment, as shown in **Supporting Figure 10**.

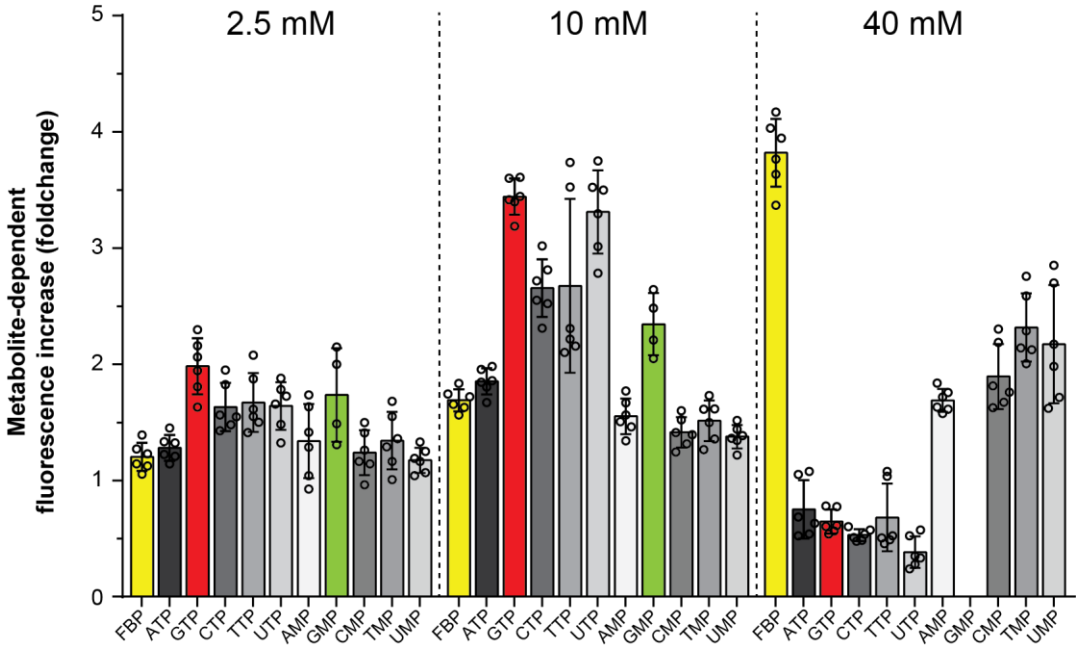


Figure 33 | S81 exhibits loose specificity to a panel of triphosphate and monophosphate nucleotides

The fluorescence emission of s81 was measured at the indicated concentrations of a panel of triphosphate nucleotides (ATP, GTP, CTP, TTP, UTP) and monophosphate nucleotides (AMP, GMP, CMP, TMP, UMP). S81 response to FBP (yellow), GTP (red), and GMP (green) was color-coded to be highlighted. The raw fluorescence values for s81 were normalized to the unspecific fluorescence variation of aptamer D1 with the different ligands. The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9, 50 mM KCl, 30 mM $MgCl_2$.

Both assays were repeated in parallel to correlate the flow cytometry results, which showed a positive fluorescence response to oligomycin A for s81 (**Figure 30**), with the intracellular metabolite level of FBP. However, intracellular FBP levels were measured by a commercially available enzymatic kit, purchased by Biovision Inc. (section 6.4) to cross-validate the results obtained by LC-MS by an orthogonal quantification method.

In the head-to-head comparison, the 4PM and Broccoli signal were used as an internal control of the flow cytometry assay and showed no response to oligomycin A (**Figure 32 a**) as in the previous assay (**Figure 30**). For the

colorimetric assay, a spiked control at increasing concentration of FBP (0.01 to 1 mM, **Figure 32 b**) confirmed the linearity range and sensitivity of the assay. As anticipated, two assays had opposite results: while s81 positively responded to oligomycin A (p-value: 0.057, **Figure 32 a**), the intracellular FBP strongly decreased (p-value: 0.004, **Figure 32 b**), thereby confirming the LC-MS results (**Figure 31**).

Taken together, the decrease in intracellular FBP upon oligomycin A treatment was confirmed by two independent methods. These results are in sharp contrast to the published findings of Rabinowitz and co-workers[15,157], in which FBP was considered a global predictor of glycolytic flux in bacteria, yeast and mammalian cells. The fact that they investigated the correlation in another cell line might justify the different result, but also indicated that FBP cannot be considered a generalized predictor of glycolysis. Given these results, the oligomycin-treatment was not valid to test s81 capability to image FBP fluctuation in cells.

On the other hand, the increase in the signal obtained by flow cytometry for 81 and not the controls (**Figure 32 a**), possibly indicated that the sensor could image different metabolites beyond the glycolytic intermediates Glu6P, FMP, and FBP, given that they all sharply decreased with oligomycin A (LC-MS for FBP in **Figure 31**, Glu6P and FMP in **Supporting Figure 10**). If the sensor could recognize other metabolites beside the glycolytic intermediates, the final readout and fluorescence signal would not be reliable. For this reason, the specificity of 81 had to be re-investigated before its applicability in cells as a metabolite-responsive sensor.

From literature, one molecule known to increase upon oligomycin treatment was adenosine monophosphate (AMP) because oligomycin causes a robust increase in endogenous AMPK activity[179]. This evidence was also confirmed by LC-MS, where AMP levels in HEK293T treated with oligomycin A was 3-fold higher than the DMSO vehicle (**Supporting Figure 11**). Given this evidence, the specificity of s81 was investigated for an exhaustive panel of nucleotides, both as triphosphate and

monophosphate, to map which target ligand could be recognized by the sensor and would cause a signal increase in the flow cytometry assay (**Figure 31**)

The specificity assay is shown in **Figure 33** and was obtained by normalizing the fluorescence results for s81 to the variation of Baby Spinach (D1) control for each condition, obtaining the metabolite-dependent fluorescence increase for s81 to the ligands (preliminary data in **Supporting Figure 12**). S81 ratiometric response to FBP overlapped to the previously performed specificity assay (**Figure 23**), albeit slightly lower at high metabolite concentration (1.2-, 1.7- and 3.8-fold at increasing FBP). However, the most surprising result was obtained at the low-millimolar level for the panel of triphosphate and monophosphate nucleotides. Undeniably, s81 exhibited a loose specificity for most triphosphate nucleotides while responding mostly to GMP between all the monophosphate ligands. The response to GTP, in particular, was robust up until 10 mM (2.0- and 3.4-fold at increasing GTP), and the same was true for GMP (1.7- and 2.3-fold at increasing GMP). At 2.5 mM, s81 exhibited a fluorescence response, which was 66% and 41% higher for GTP and GMP, respectively, than FBP. However, s81 response plummeted at 40 mM GTP, whereas the response at 40 mM GMP could not be obtained due to the visible precipitation of the metabolite at this concentration.

Given these results, it was concluded that s81 was specific to FBP only at a high concentration of the metabolite, which was also true compared to other glycolytic metabolites (**Figure 23**). At lower concentrations, the vast majority of tested analytes exhibited some response, albeit at a different level. A pattern of response for s81 was noted when parallelizing the specificity assays performed for the glycolytic intermediates Glu, Glu6P, FMP, and FBP (**Figure 23**) with the triphosphate (NTPs) and monophosphate nucleotides (NMPs) (**Figure 33**). Specifically, the highest response with s81 seemed to depend on the level of phosphorylation of the given molecule: the highest response at the lowest analyte concentration was obtained with NTPs, then the bi-phosphorylated FBP, and finally the lowest response was obtained with NMPs. Moreover, glucose elicited no

response at any concentration. Given this pattern, it can be speculated that the most critical feature of a tester analyte that will elicit s81's response is the phosphate group. The higher the phosphorylation, the stronger seems to be the response of s81. On the other hand, the metabolite GMP stands out of this picture, given its robust capability to elicit s81's response, which surpassed even FBP at ≤10 mM. In that sense, it could be said that s81 is much rather a sensor for GMP than FBP.

To conclude, s81 undeniably showed a ratiometric response to FBP in the wide millimolar range, but its promiscuity towards other glycolytic intermediated and even most tri- and monophosphorylated nucleotides indicated that its applicability as a sensor for any given target would be limited, especially in the cellular context.

4 DISCUSSION

This work explored the concept of *in vitro* selection of functional RNA aptamers and ventured it towards a novel and worthwhile goal, that of direct identification of allosteric RNAs, a class of bispecific RNA comprised of two analyte-binding molecules allosterically interlinked to one another. Specifically, this study aimed to generate a variant of Baby Spinach aptamer, a fluorogenic RNA affine to the non-fluorogenic dye DFHBI[94], which responded ratiometrically to fructose 1,6-bisphosphate (FBP). FBP is a metabolic intermediate of glycolysis with a central role in glycolytic homeostasis and regulation (Section 1.1.4).

Initially, the identification of FBP-responsive Spinach variants was investigated by SELEX, an *in vitro* selection scheme based on affinity separation (Section 1.2.1). A plethora of SELEX schemes were developed to obtain aptamers with certain specifications[180], but there are no examples of its capability to enrich RNA aptamers where the binding of a small-molecule reporter, such as DFHBI, is allosterically-triggered by an orthogonal target molecule. On the other hand, SELEX was successfully applied to isolate aptamers with strand-displacement capability from a library immobilized to a complementary strand[138] (Section 1.3.1). Different SELEX methods were developed to isolate structure-switching DNA aptamers

directly[181,182], and recently library-capturing was also applied for RNAs[140]. These examples indicate that binding-induced structural change caused by the interaction with a small molecule, e.g., FBP, could be employed to drive the selection pressure.

A traditional SELEX protocol was first established to enrich and identify fluorogenic RNA of DFHBI, following the published results of Jaffrey and co-workers[77]. The protocol was reproduced successfully and re-identified Spinach aptamer, albeit in its truncated version Baby Spinach (**Figure 5- Figure 7**). However, the attempts to identify an improved variant of Spinach specific for a red-shifted version of DFHBI (**Figure 8**) or folded in intracellular-mimicking buffer system (**Figure 9**) did not succeed, likely due to sub-optimal selection conditions. For instance, the synthesis of the red-shifted DFHBI was not stereospecific, and in the intracellular-mimicking phosphate buffer, the RNA phosphate charge is less neutralized.

Consequently, the SELEX protocol for selecting fluorogenic RNA of DFHBI was repurposed to select allosteric RNAs that bind DFHBI in a target-dependently (**Figure 10**). The newly-designed allosteric SELEX protocol hinged on depleting the canonic DFHBI-binders via an initial negative selection followed by a positive selection step designed to retrieve target-dependent binders. An incubation step was added between these two steps to promote the interaction between RNAs and the target of choice.

A panel of small-molecule (acetyl-coenzyme A), microRNA (miR21), and protein (lysozyme) was employed in parallel to FBP to maximize the chemical diversity from the target and thus increase the chance of one successful selection. Yet, the enriched target sequences within the selection cycles could not be tested using the traditional post-SELEX characterization methods relying on radioactive-[161] or fluorescent-labeled RNA[61]. These methods were not suitable to monitor the enrichment of fluorogenic sequences that complex DFHBI target-dependently. Instead, the increase in DFHBI-fluorescence of an enriched pool in the presence of

the target was used as a proxy for the enrichment of sequences of interest. Despite the panel of target ligands, the allosteric SELEX protocol ever failed to enrich allosteric RNAs of DFHBI from a stem-loop primed random library (**Figure 11**).

When a generalized library fails, the introduction of structured modules within the random region is a widely applied methodology to increase the selection success. Szostak and co-workers were the first to apply this concept by identifying high-affinity GTP aptamers in a stem-loop primed library and not in a random one[142]. Spinach is another case, since Jaffrey and co-workers identified this aptamer only in Szostak's steam-loop primed library[77]. Later, Szostak analyzed the reason behind the improved performance of structured libraries introducing the functional-information content of a library, which is defined as the probability to select a given aptamer from a given library[183]. As an example, Szostak stated that the probability of identifying an ATP-binding aptamer (with a K_D in the micromolar range) is 10^{-11} in an N70 library, thus elegantly expressing how the emergence of an aptamer is a function of the ensemble of the tested sequences of a library[183]. Therefore, in the stem-loop primed library, its functional-information for GTP- or DFHBI-binding aptamers is greater than for a random library.

Following this theory, and capitalizing on the wealth of structural information available for the Spinach-DFHBI pair (**Figure 2, Figure 3**), a panel of Spinach-based specialized libraries was designed. By inserting Spinach's critical nucleotides, the library's functional-information should shift the sequence space from a conventional to a specialized subset with improved interaction to DFHBI, wherefrom to enrich the ones with allosteric function over the canonic binders. On the other hand, a library too primed in DFHBI sequence would not be hindered by the allosteric protocol's negative selection step and could, consequently, prevent the selection of allosteric binders. With that in mind, a balance between random and constant sequences was found in library S2 and S6 (**Figure 12 b, g**), in which Spinach's structural stems P3 and P2, respectively, were pointedly randomized to

achieve minimal background fluorescence, but the capability to reselect Spinach in a DFHBI-SELEX protocol (**Figure 12 i, k**).

To optimize an allosteric SELEX protocol combined with the specialized libraries S2 and S6, the target adenosine 3′, 5′-cyclic monophosphate (cAMP) was picked as a test molecule due to its biological relevance[184] as well as "aptamerogenicity", likelihood generate an aptamer, being a ribonucleotide-derived small molecule[64,142,154,159]. Expectedly, equipping the libraries with Spinach's core nucleotides had the side-effect of enriching the canonic DFHBI-binders population (**Figure 13 a, c**), and the negative selection stringency was adjusted to counteract their enrichment (**Figure 13 b, d**). Library S2 exhibited no fluorescence response to cAMP throughout the selection (**Figure 13 b**), while S6 promisingly exhibited a minimal but significant cAMP-dependent fluorescence increase (**Figure 13 d**). However, the attempts to obtain a robust increase in the cAMP signal by increasing the negative selection steps (**Figure 13 e**) or restarting selection from an earlier selection cycle (**Figure 13 f**) were unsuccessful. The previously tested targets miR21, lysozyme, fructose 1,6-bisphosphate, and acetyl-coenzyme A (**Figure 14**) were tested in combination with S6 but with no positive outcome.

To conclude, optimizing the available information of a library is a worthwhile strategy to identify novel designs with emerging properties. For instance, inserting the sequence of Spinach within the random region had a direct and measurable effect on the library's background affinity towards DFHBI. The results could also be rationalized and put into context with the randomization of Spinach paternal aptamers. Another remarkable result was the quantifiable response of the specialized libraries against the negative selection, which was useful to fine-tune the stringency from cycle to cycle. Albeit no allosteric RNAs were identified, the specialized libraries offered a window to study the impact that negative selection had on the sequence population without interrupting the SELEX protocol, which other methods such as NGS would require. With this method, the influence of the negative selection could be investigated and generalized to other selections where

this step was known to harbor a positive effect on the final aptamer emergence, such as for the theophylline[185] or vascular smooth muscle-internalizing aptamers[146].

In any case, the generalized approach to identify allosteric binders of DFHBI for a target of choice was not successful. Therefore, another strategy was investigated to generate an FBP sensor by emulating the rational designs of conjoined aptamer modules (section 1.2.3). Specifically, the generation of an allosteric RNA sensor was explored by conjoining the fluorogenic RNA Baby Spinach[94] with aptamer C45[161], which binds FBP in the millimolar range[162] (**Figure 15**).

FBP is a key metabolite consistent with glycolytic flux in bacteria[47,48], yeast[44-46], and mammalian cells[15,157], and molecular tools capable of probing this metabolite in real-time are highly demanded. In particular, the development of a millimolar-range sensor for FBP could have immediate applicability in single-yeast cells metabolomics imaging[46] and, in combination with glycolytic boosters, could be applied in mammalian cells to replace current methods that measure glycolysis with population-averaging approaches (Section 1.1.3) and even allow to investigate the metabolic heterogeneity within a cellular population (Section 1.1.2).

Rational designs are limited in the throughput of screenable aptamers combinations, which usually spans from 5 to 10[122-126,128,129,131,132] (**Table 1**). To circumvent the need to test each construct, the structural information of C45[161] and Baby Spinach[94] aptamers were embedded into a structure-guided chimeric RNA library to identify the best allosteric designs (**Figure 15**). So designed, the chimeric library consisted of ~4 million different sequences in which Spinach's core domain was unfolded in the absence of FBP, and a selection method was applied to identify those iterations which would refold Spinach's core target-dependently.

In combination with the allosteric SELEX protocol, the chimeric library exhibited a promising increase of FBP-dependent fluorescence (**Figure 16 a**) that was not obtained with the stem-loop (**Figure 11 b**) and P6 libraries (**Figure 14 c**). On the other hand, this approach yielded truncated by-products[164] that ended up dominating the library pool, as apparent by the agarose gel analysis of the pool

progression (**Figure 16 b**) and NGS analysis (**Table 3**). The NGS analysis also revealed randomization of the sequence information within the chimeric library (**Supporting Figure 3**), which is an alarming sign that the selection did not progress as expected. Most likely, the enrichment of by-products with randomly truncated sequence information scattered the chimeric library's structural identity. Likewise, the top-NGS candidates (**Table 3**) showed a sequence trimming to ~ 50 nt RNA length from the original 118 nt of the library and retained no fluorogenic properties (**Figure 16 c**). The by-product could not be removed from the pool by conventional purification methods, and, thus, no additional selection cycles were performed.

Throughout all performed selections, the allosteric SELEX protocol failed to enrich sequences with allosteric binding properties. Nevertheless, the FBP-dependent response obtained in the allosteric SELEX binding assay (**Figure 16 a**) might indicate that allosteric sequences are within the chimeric library. To specifically scout them, a method with a more substantial power of discrimination was required, but that is based on fluorescence screening instead of sequence retention on affinity matrix[116]. The microfluidic-assisted *in vitro* compartmentalization (µIVC) pipeline engineered by Ryckelynck and co-workers was employed to screen single RNA species of a pool according to their fluorescence capacity[114,163]. This µIVC pipeline already demonstrated its proficiency to identify fluorogenic RNAs and even allostery-regulated ones (Section 1.3.2). The µIVC method, combined with an ad hoc designed screening protocol, allowed the isolation of a fluorogenic RNA pool responsive to FBP (**Figure 17**).

Similarly, the allosteric SELEX protocol enriched an FBP-responsive fluorescence signal, but randomization of the structural information within the chimeric library was revealed by NGS (**Supporting Figure 3**), together with a lack of selection stringency against non-target sequences (**Figure 17 d**). Contrarily, the µIVC pipeline maintained the structural information of the chimeric library (**Supporting**

Figure 5) while sharply de-enriching the pool of non-target sequences (**Figure 17 d**).

The sequencing of the µIVC-screened RNA pool revealed a co-enriched family of truncated RNA molecules that lacked the ~13 nt region designed to hinder the folding of Baby Spinach's G-quartet (**Supporting Table 1, Supporting Table 2**). Due to the lack of this critical region, this family lost, on average, its capacity to interfere with DFHBI-binding ability (**Figure 18 a**). Parallelly to the truncated family, full-length RNA sequences were enriched and identified, although no sequence revealed high copy numbers (0.75-0.40% for full-length RNAs, **Supporting Table 1**).

The persistence of full-length sequences, together with a reduced degeneration of the constant sequences compared to the allosteric SELEX protocol (**Table 3**), points at the contribution of µIVC screening in protecting the system from unwanted events, e. g., recombination during PCR and parasite dissemination[164]. While these results cannot be yet generalized, the comparative NGS data indicates that µIVC pipelines based on fluorescence-based screening out-performed the affinity matrix-based SELEX approach in selecting and identifying FBP-responsive fluorescent RNA sequences.

A sensor candidate panel was tested by gathering sequences from the top-enriched NGS hits (**Supporting Table 1**) and SANGER sequences clones (**Supporting Table 2**). While all sensor candidates responded to FBP somewhat, s81 exhibited the most robust fluorescence light-up to FBP both in selection buffer and minimal fluorescence buffer (**Figure 18**). S81 was chosen as an RNA-based fluorescence sensor for FBP and further characterized.

Initially, s81 was identified by SANGER sequencing, and only a subsequent deep-sequencing analysis allowed the identification of this clone within the full-length NGS pool (**Supporting Table 3, Supporting Table 4**). The finding that s81 is an outcast of selection, whereas the shortened by-products are the most enriched by sequence analysis, demonstrated the current limitations of the applied screening

protocol. While some troubleshooting is due, the µIVC pipeline is a promising pipeline to generate RNA-based fluorescence sensors from pre-conceived aptamers and out-performs the rational design-build-test alternatives by bringing the throughput from ca. 10 to more than 10^6 iterations screened.

Next, the performance of s81 as an FBP-responsive fluorescence sensor was characterized *in vitro*. S81 responded ratiometrically to FBP in the range of 2-600 mM, but its fluorescence signal strongly depends on Mg^{2+}-ions (**Figure 19**). Mg^{2+}-ions are required by both paternal aptamers – Baby Spinach and C45 – that constitute s81[94,162], but their levels are influenced by FBP, capable of chelating divalent cations *in vitro*[167]. Therefore, s81 functions as a sensor forasmuch as Mg^{2+}-ions are in excess to FBP. However, this FBP- Mg^{2+} balance is only relevant *in vitro* because FBP chelates Fe^{2+}-ions[167] in cells, and the concentration of cellular Mg^{2+}-ions has previously allowed the application of Spinach-based fluorescence constructs (Section 1.2.2). Likewise, the C45-domain of 81 should also fold in cells, as it exhibits a similar Mg^{2+} requirement to Spinach[154].

Conversely, increasing the concentration of K^+-ions *in vitro* resulted in an enhanced baseline fluorescence of 81 and a proportional decrease of the FBP-fluorescence response (**Figure 20**)., This is most likely caused by the capability of K^+-ions to stabilize Baby Spinach's G-quadruplex[104] and indicates that the allosteric control of s81 is based on the G-quadruplex destabilization, as expressly designed into the C45-Baby Spinach chimeric library. Another evidence of the FBP-dependent G-quadruplex destabilization by s81 was the point mutation analysis of s81's CACA stretch (**Figure 21**). This CACA stretch was designed explicitly within the chimeric library to interact with Baby Spinach's G-core (Section 3.4.1). The progressive mutation of the CACA-stretch induced a gradual increase of the baseline-fluorescence and curbed the FBP-fluorescence response of s81, highlighting the centrality of the CACA stretch for the sensor's allosteric control. Furthermore, a specific point-mutant control was identified (4PM, **Figure 21**) in which FBP-dependency on fluorescence is suppressed.

In an optimized ion buffer at 50 mM KCl and 200 mM $MgCl_2$, s81 reached the maximal sensing range up to 250 mM FBP (**Figure 19**). In this milieu, s81 exhibited no interaction to DFHBI without FBP, whereas an EC_{50} value of 5.56 ± 0.40 µM was measured in the presence of 250 mM FBP (**Figure 22**). Nevertheless, the EC_{50} value was ~4.9-fold less affine than D1 aptamer control (1.13 ± 0.07 µM), indicating that s81 reduced the affinity to DFHBI while concomitantly gaining allosteric control towards FBP.

As for sensing specificity, s81 showed to interact with Fructose 1,6-bisphosphate (FBP) mainly compared to a panel of structurally similar metabolites (**Figure 23**). Nevertheless, 81 exhibited cross-specificity for Glucose 6-phosphate (G6P) and Fructose 6 phosphate (FMP); these metabolites are precursors of FBP in glycolysis (Section 1.1.1).

For its application as a cellular FBP sensor, s81 was transiently expressed in mammalian cells using the Tornado expression vector[135], upon verification of the functional integration (**Figure 27**). However, the cellular fluorescence intensity of s81 matched the *in vitro* fluorescence only in HEK293T (**Figure 28**), an engineered cell-line that allows high copy-number RNA expression[186].

The insufficient fluorescence output in cells is a known flaw of the Spinach-DFHBI fluorogenic system, which could only be circumvented with optimized dye variants[109], extensive RNA scaffolding[112], and cellular expression as circular RNA[135]. Despite these improvements, fluorogenic constructs based on the Spinach aptamer seem to perform well only in the HEK293T line, raising questions about the applicability of Spinach as a generalized cellular tool. On that note, Bouhedda *et al.* compared the currently available fluorogenic RNA aptamers in Hela cells and confirmed that many lack signal robustness and photostability[98]. Instead, they proposed a novel RNA-dye concept based on self-quenched sulforhodamine B dyes that could prove to be a valuable alternative.

Nonetheless, HEK293T was identified as a suitable cellular system for s81 expression and to test the performance of s81 in probing intracellular FBP variation.

At large, the physiological concentration of FBP in mammals is ≤1mM[172,173] and is too low to be detected by s81. Advantageously, Rabinowitz and co-workers demonstrated that intracellular FBP levels robustly increased by oligomycin A-treatment[15,157], an antibiotic widely used to measure the maximum glycolytic capacity of cultured cells[33,175,176]. Oligomycin inhibits the mitochondrial complex V[187] and compels shunting of the ATP-production burden from mitochondria's respiration to cytosolic glycolysis (**Figure 29**), leading to a substantial boost in the glycolytic flux and, supposedly, intracellular FBP. Encouragingly, the oligomycin-induced glycolytic boost caused a modest but significant shift in the fluorescence of s81 expressing HEK293T cells (**Figure 30**). In contrast, no difference was reported by 4PM and Broccoli controls, demonstrating that the variation was due to an allosteric rearrangement of the sensor and not on other confounding factors, such as RNA expression or unspecific fluorescence fluctuation.

Contrary to expectation, FBP decreased upon oligomycin A-treatment by LC-MS quantification of the metabolite extracts (**Figure 31**). A corresponding result was obtained with G6P and FMP, the glycolytic intermediated for which s81 exhibited cross-specificity (**Supporting Figure 10**). Further, the discord between the s81 signal response and FBP signal decrease was confirmed by head-to-head analysis using an orthogonal quantification method (**Figure 32**). These results indicated that s81 could sense another one, or more, cellular metabolites beside FBP, FMP, and G6P, triggering a response with oligomycin A. In the attempt to identify the cellular target, an *in vitro* assessment of s81 specificity evidenced that the sensor unspecifically recognized different phosphorylated nucleotides, with varying fluorescence response (**Figure 33**).

This *pan*-specificity of s81 is a critical negative feature that hampers the applicability of this sequence as a sensor, and is in opposition with the specificity exhibited by the paternal C45 aptamer, which exhibited specific elution from an FBP-functionalized resin by FBP and not G6P, FMP, or ATP at 10 mM[162]. At the same analyte concentration, s81 is only 50% less specific to G6P and FMP

compared to the target FBP (**Figure 23**), whereas ATP was even more specific than all glycolytic intermediate (**Figure 33**). The truncation and randomization of C45 used to design the chimeric library could be the cause for the decrease in the specificity of the selected sensor s81.

However, the low-enrichment profile of s81 by NGS (**Supporting Table 4**) might indicate that a much more specific variant, albeit with a lesser fluorescence response, is present amongst the previously tested candidates (**Figure 18**) or the newly identified sequences, such as sequence A1 (**Supporting Table 3**). Therefore, a fundamental step for future research is to screen the panel of sensor candidates first for specificity and only secondly for performance, in order to identify a better FBP-responsive variant.

Another promising area for future research is to capitalize on the degenerated specificity of s81 towards mono- and triphosphate nucleotides, GMP and GTP in particular (**Figure 33**), to identify sensors specific for these metabolites. Specifically, the loop region of the chimeric library originally belonging to C45 (G26-C36, **Figure 15 a**) could be randomized to create a 12N functional library from which to screen, in combination with the µIVC pipeline, responsive fluorogenic RNAs for these metabolites.

However, the combination of specialized fluorogenic library designs and ultra-HT µIVC screening could be applied beyond the pure fluorogenic RNAs field. For example, allosteric variants could be identified from the scaffolds of well-established ribozymes to engineer small-molecule responsive RNA switches. In combination with AAV-based delivery, these constructs could have a tremendous impact in the novel field of ribozyme therapy[188,189].

5 METHODS

5.1 Preparation of the affinity matrix

DFHBI-Sepharose was prepared following the protocol from Paige et al.[77] Amine-functionalized DFHBI was first dissolved in 1.2 mM HCl to facilitate solubilization and diluted in 50 mM Borate buffer pH 7.5 at the concentration of 1 mM. This fluorophore solution was added to NHS-activated Sepharose, which was previously equilibrated with 3 volumes of ice-cold borate buffer. In proportion, 1 nMol of the fluorophore was added every mg of Sepharose. The resin was incubated with fluorophore solution rotating overnight at 4 °C in the dark. The resin was washed with borate buffer and incubated with 1 M methylamine 50 mM Borate buffer pH 7.5 for 3 h at 4 °C to react with any remaining NHS-activated sites. After thorough washing with alternating 50 mM KAc pH 5.0 and 100 mM boric acid pH 8.0, the resin was stored in 1:1 EtOH:100 mM NaOAc pH 5.4 at 4 °C. The efficiency of Sepharose coupling was monitored by measuring the absorbance at 400 nM of free fluorophore in the flow-through and was estimated to be >90%. Mock Sepharose was prepared following the same protocol, but the over-night coupling solution was not supplemented with amine-functionalized DFHBI.

5.2 RNA Purification for library pool and specific sequences

Double-stranded DNA of the library pool for *in vitro* transcription was obtained from the single-stranded DNA templates following a large-scale amplification protocol, which uses a gentle condition to minimize over-amplification.

For Large-scale PCR condition, 1 nMol (~ 10^{15} different sequences) of ssDNA was used in a 10 mL PCR reaction using GoTaq DNA polymerase enzyme supplied with 0.25 mM dNTPs, 1 µM primers, 3 mM $MgCl_2$ and was amplified (2 min 95°C; 30 sec 95°C, 30 sec 58 °C, 1 min 72°C; hold 4°C) for 6-8 cycles depending on the library. The forward primer with a sequence complementary to the constant region was used for amplification and to introduce a T7 promoter sequence (5'-TAATACGACTCACTATA -3'). The final dsDNA product was purified by MicroSpin G-25 columns to remove nucleotides, Sodium acetate-EtOH precipitated, and resuspended in ddH_2O. Based on the PCR cycling and efficiency, ~1 nMol of dsDNA was transcribed overnight in a 1 mL *in vitro* transcription reaction containing 2.5 mM NTPs, 25 mM $MgCl_2$, 5 mM DTT, 0.4 U/µL RNasin and 0.5 U/µL T7 polymerase in 40 mM Tris-HCl pH 7.9 buffer.

Similarly, the dsDNA amplification for a specific sequence was prepared with the same polymerase system, but using the manufacture's PCR instruction. While the PCR cocktail was the same, the template concentration was set at 0.01 µM and the cycling (2 min 95°C; 15 sec 95°C, 15 sec 58 °C, 30 sec 72°C; hold 4°C) adjusted to obtain sufficient product, typically 12-16 cycles. Then, the PCR product was precipitated with Sodium acetate-EtOH, and the pellet resuspended in ddH_2O to be used for *in vitro* transcription. As a rule of thumb, 100 µL of PCR reaction were resuspended after precipitation in a final volume of 100 µL T7 reaction to obtain enough RNA for downstream assays. All transcriptions were incubated at 37°C overnight.

Next, reactions were treated with RNAse-free DNAse I according to the manufacturer's protocol; the transcribed RNA was purified by Phenol/Chloroform

extraction and Sodium acetate-EtOH precipitation. The RNA pellet was resuspended in 8 M Urea and loaded on a 10% TBE-Urea acrylamide gel for gel purification. Full-length RNA was visualized and cut out of the gel with UV shadowing and recovered from the gel with a crush-soak method, precipitated and resuspended in ddH_2O for quantification with NanoDrop™ 2000c (Thermo scientific).

5.3 *In vitro* fluorescence measurements

In all *in vitro* fluorescence assays, the RNA – either library or synthesized single sequence – was purified by Urea-PAGE gel followed by crush-soak method and was fixed at the concentration of 1 µM. The fluorophore used was DFHBI at the concentration of 10 µM, while it was substituted for its enhanced variant DFHBI-1T for screening and cellular application. The end-point fluorescence value was used as a proxy of the formation of the RNA-fluorophore complex.

The initial sensor candidate shortlisting was performed in selection medium (2 mM NTP, 25 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, 1 mM Spermidine, 0.1% of Pluronic F68, 1 µL of PPase in 40 mM Tris-HCl pH 7.9), while all the other assays were performed in minimal fluorescence buffer (50 mM KCl in 40 mM Tris-HCl pH 7.9) with a concentration of $MgCl_2$ dependent on the concentration of FBP in that assay. For EC_{50} determination, the affinity curves in the presence and absence of FBP were determined using nonlinear regression analysis in Prism software and matched by least-squares fitting to a standard dose-response model for 1:1 complexation. Values represent the mean of three independent experiments ± 1 standard deviation. Fluorescence emission was recorded using EnSpire Multimode Plate Reader (PerkinElmer) with the following instrument parameter: excitation wavelength, 462 nM; emission wavelength, 503 nM; slit widths, 1 nM.

5.4 SELEX protocols

5.4.1 Test SELEX for DFHBI

500 pmol RNA produced from the stem-loop library S1 was resuspended in 100 µL selection buffer (125 mM KCl, in 40 mM HEPES pH 7.4). The RNA was heat-denatured at 80 °C for 5 min and snap-cooled down to 30°C to promote RNA folding. $MgCl_2$ was added to the solution at the concentration of 5 mM. The affinity matrix for the negative selection step was prepared by incubating 15 mg of Mock sepharose with 100 µL of 5 µM tRNA in the selection buffer. After 15 minutes pre-incubation with the tRNA solution, the RNA solution was added to the mix and incubated for 30 minutes at 30 °C. The flow-through was collected and incubated for 30 minutes with 15 mg DFHBI-Sepharose as a positive selection step. The affinity matrix was washed 3 times with 100 µL selection buffer, and bound RNA was finally eluted at 80 °C per 20 minutes in 40 µL H2O. The eluted RNA was reverse-transcribed using eAMV RT following the manufacturer's protocol, followed up by PCR amplification and T7 transcription over-night – same protocols as for the library preparation – to yield the pool for the next round of selection. To visualized the enrichment of shorted parasitic sequences, transcribed RNA was resolved on 4% agarose Gel with Agarose LE pre-stained with Ethidium Bromide. As shorter RNA sequences were visualized, the RNA library was gel purified following the same protocol as for library preparation. Conversely, the T7 RNA pool was used for the next step of selection. After each selection round, the RNA input was gradually decreased (from 500 to 50 pmol) together with the affinity matrix (from 15 to 5 mg) to increase the stringency and permit the enrichment of high-affinity aptamers.

5.4.2 Selection of Spinach aptamer improved variants

The SELEX protocols for selecting improved variants of Spinach were designed accordingly to the test SELEX for DFHBI (Section 5.4.1) with some essential adjustments.

The initial eDFHBI SELEX (**Figure 9 a**) was conducted on the stem-loop primed library using eDFHBI-charged NHS-activated Sepharose; this matrix was obtained as in section 5.1, but with amine-functionalized eDFHBI[150].

The second eDFHBI SELEX (**Figure 9 c**) was conducted on the S0 library at an initial amount of 50 pmol, using the eDFHBI-charged matrix. The RNA amount was not changed throughout the selection, whereas the affinity matrix was gradually decreased from 15 to 5 mg and maintained constant from cycle 3 onwards.

The third SELEX (**Figure 9 d**) was conducted on the intracellular mimicking buffer (160 mM KCl, 5mM NaCl, 1 mM $MgCl_2$ in Phosphate buffer pH 7.2).

5.4.3 Allosteric SELEX on the stem-loop primed library

For all selections presented in **Figure 11**, 500 pmol RNA from library S1 was resuspended in 100 µL selection buffer (125 mM KCl, 40 mM HEPES pH 7.4).

The RNA was heat-denatured at 80 °C for 5 min and snap-cooled down to 30°C to promote RNA folding. $MgCl_2$ was added to the solution at the concentration of 5 mM. The affinity matrix for the negative selection step was prepared by incubating 15 mg of DFHBI sepharose with 100 µL of 5 µL tRNA solution in the selection buffer. After 15 minutes of pre-incubation with tRNA, the RNA solution was added to the mix and incubated for 30 minutes at 30 °C. This negative selection step was crucial for depleting the population in target-independent DFHBI binders and was repeated once for each selection cycle. The flow-through was collected and incubated for 15 minutes with the target of interest – 200 µM AcCoA for SELEX **a**, 5 mM FBP for SELEX **b**, 1 µM miR21 for SELEX **c,** and 0.2 µM lysozyme **d** – to allow the structure-switching to occur. The solution containing a pre-enriched RNA library and the choice target was incubated again for 30 minutes with 15 mg DFHBI-Sepharose as a positive selection step. The affinity matrix was washed 3 times with 100 µL selection buffer supplemented according to target concentration. Finally, the bound RNA was eluted at 60 °C per 30 minutes in 40 µL

H_2O. The eluted RNA was treated accordingly to Section 5.4.1 to purify the input RNA for the next selection cycle.

5.4.4 Specialized library designs: test SELEX and allosteric SELEX

The DFHBI SELEX presented in **Figure 12 i-k** were conducted according to Section 5.4.1 on the specialized libraries S2 (SELEX **i**), S3 (SELEX **j**), and S6 (SELEX **k**).

The allosteric cAMP SELEX presented in **Figure 13 a, c** were conducted according to Section 5.4.3 on the specialized libraries S2 (SELEX **a**) and S6 (SELEX **c**) using 1 mM concentration of the cAMP target.

The allosteric cAMP SELEX presented in **Figure 13 b, d** was conducted according to Section 5.4.3 on the doped libraries S2 (SELEX **b**) and S6 (SELEX **d**) using 3x negative selections steps for each cycle and 1 mM concentration of the cAMP target.

The allosteric SELEX presented in **Figure 14** was conducted according to Section 5.4.3 on the doped library S6 using a negative selection step, which was repeated once in cycle 1, twice in cycle 2-3, thrice from 3-onwards.

5.4.5 FBP Allosteric SELEX on chimeric library

The SELEX presented in **Figure 16** was conducted on the chimeric library (500 pmol input RNA), which was resuspended in 100 µL selection buffer (50 mM KCl in 40 mM Tris-HCl pH 7.9). The RNA was heat-denatured at 80 °C for 5 min and snap-cooled down to 30°C to promote RNA folding. $MgCl_2$ was added to the solution at a concentration of 25 mM.

The affinity matrix for the negative selection step was prepared by incubating 15 mg of DFHBI sepharose with 100 µL of 5 µL tRNA solution in the selection buffer. After 15 minutes of pre-incubation with tRNA, the RNA solution was added to the mix and incubated for 30 minutes at 30 °C. This negative selection step was crucial

for depleting canonic DFHBI binders and enriching FBP-dependent binders. It was repeated once in cycle 1, twice in cycle 2-3, and thrice from 3-onwards. The flow-through was collected and incubated for 15 minutes with 20 mM FBP at 30 °C to allow structure switching. The solution containing the pre-enriched RNA library and FBP were incubated again for 30 minutes with 15 mg DFHBI-Sepharose as a positive selection step. The affinity matrix was washed 3 times with 100 µL selection buffer supplemented with 20 mM FBP, and bound RNA was finally eluted at 60 °C per 30 minutes in 40 µL H_2O.

The eluted RNA was reverse-transcribed using eAMV RT following the manufacturer's protocol, followed up by PCR amplification and T7 transcription O/N – same protocols as for the library preparation – to yield the pool for the next round of selection. Finally, the bound RNA was eluted at 60 °C per 30 minutes in 40 µL H_2O. The eluted RNA was treated accordingly to Section 5.4.1 to purify the input RNA for the next selection cycle.

5.5 µIVC screening

The microfluidic chip development and the droplet digital PCR preparation followed the protocol from Autour *et al.*[114].

However, the droplet occupancy was fine-tuned in the first screening cycle to accommodate a higher variability, increasing the initial droplet occupancy of the ddPCR to an occupancy of 65 % (λ=1) and screening for more than 5 million droplets. Occupancy was decreased in cycles 2 and 3, accounting for 17% and 27%, respectively. The ddPCR emulsion (made of 2.5 pL droplets) was PCR-amplified off-chip. Then, each thermocycled PCR droplet was fused 1 to 1 with a droplet containing an *in vitro* transcription (IVT) mixture supplemented with FBP for a final concentration of 2.5 mM for the first two positive screening cycles, while in the third cycle FBP was omitted. The IVT mix contained 2 mM NTPs, 25 mM $MgCl_2$, 40 mM Tris-HCl pH 7.9, 50 mM KCl, 5 mM DTT, 1 mM Spermidine, 0.1% of Pluronic F68, 1 µL of PPase, 20 µM DFHBI-1T, 0.5 µM of Dextran-Texas Red, 0.5

U/µL L T7 RNA polymerase. The IVT mix was kept on ice during the droplet fusion; the emulsion was collected off-chip in a tube and incubated for 1 hr at 37°C to allow RNA transcription to occur. Droplet sorting was carried out by re-injecting the emulsion on a microfluidic sorting device designed to analyze the green (DFHBI-1T) and orange (Texas Red) fluorescence of each droplet.

Cycles 1 and 2 of positive screening were performed in the presence of 2.5 mM FBP, and the top greenest droplets – 0.02 and 0.46 %, respectively – displaying an orange fluorescence corresponding to single fused droplets were gated. Next, cycle 3 negative screening was performed in the absence of FBP, and gates screened for nil (4%), dim (4%), and highly fluorogenic (2.6%) droplets with an orange fluorescence corresponding to single fused droplets. In both screening conditions, gated droplets were deflected into the collection channel by applying an AC field (1000 V 30 kHz) and collected off-chip. Finally, sorted droplets were treated with 100 µL 1H, 1H, 2H, 2H-perfluoro-1-octanol to break the emulsion by vortexing, and the DNA-containing aqueous phase was retrieved.

5.6 NGS

Next-generation sequencing (NGS) analysis was performed on NextSeq500 with High Output v2 chemistry (Illumina) following Tolle *et al.*[190]. First, the chosen samples were re-amplified using indexed primers and purified by Nucleospin PCR columns. 0.125 µg of purified DNA for each sample was used for ligation of the adapter needed for immobilization and processing of the sample on the NextSeq500 using TruSeq DNA PCR-Free sample preparation kit.

Ligation was performed according to manufacturer instructions. Afterward, the ligated products were resolved on a 2% agarose gel and purified using Nucleospin Gel clean-up columns, following manufacturer instructions. DNA was resuspended using kit resuspension buffer and was used for Illumina sequencing. NGS analysis was performed with the in house developed software. NextSeq 75 cycles

sequencing kit was used for the analysis of all rounds. Due to the incapability for the sequencing kit to cover the full library sequence space, the forward (FW) and reverse (RV) NGS runs were in silico analyzed separately, and the RV run. Since the RV run covers the randomized region of the library showed higher variability than the FW run, we used the RV run for top-enriched sequence identification and de-enrichment analysis. The FW run was used to obtain the full sequence of the top-enriched sequences exploiting the overlapping sequenced region between the two runs.

5.6.1 Specs test DFHBI SELEX analysis

For the test DFHBI SELEX, 40 sequences with >0.1% enrichment were identified in silico in the RV run (69.3 %-0.1%) and FW run (69.3 %-0.1%) of cycle 5 and were organized in datasets with selection cycles 0, 2 and 4. The top-5 sequences of the RV dataset (67.31 %-0.56 %) were aligned to the top-enriched FW sequences to obtain the full-length enriched sequences (**Table 2**).

	Index	**Sequence count**	**Different sequences**	**Sequence variability**
Starting pool	TGACCA	10261	10166	99%
Cycle 2	TTAGGC	64688	62639	96%
Cycle 4	CGATGT	88845	58420	67%
Cycle 5	ATCACG	1516573	145452	9%

Table 4 |NGS indices and sequence representation for the test DFHBI SELEX cycles.

5.6.2 Specs Allosteric SELEX-µIVC screening NGS analysis

For allosteric SELEX, 81 sequences with >0.1% enrichment were identified in silico in the RV run of cycle 7 (2.18%-0.01%) and organized in datasets with selection cycles 0, 5, and 6. The top-5 sequences of this dataset (2.18%-0.42%) were

aligned to the top-enriched FW sequences to obtain the full-length enriched sequences (**Supporting Table 1**).

Since these sequences were highly truncated, we mined the RV top-enriched dataset for untruncated sequences using the constant region of C45 in the chimeric library (5'-CACANGTGTTGACGACTCGCTG-3'), but no hit was identified.

	Index	**Sequence count**	**Different sequences**	**Sequence variability**
Starting pool	CTTGTA	223781	218097	97%
Cycle 5	GGCTAC	1370817	939882	68%
Cycle 6	TAGCTT	60587	50312	64%
Cycle 7	GATCAG	856145	540286	63%

Table 5 | NGS indices and sequence representation for the allosteric SELEX on the chimeric library.

The same rationale was applied for µIVC screening, where 41 sequences with >0.1% enrichment were identified in silico in the RV run of cycle 3 gate 2 (3.28%-0.14%) and organized in datasets with selection cycles 0 and 2. The top-30 sequences of this dataset (3.28%-0.22%) were aligned to the top-enriched FW sequences to obtain the full-length enriched sequences, which were aligned to the library pool (**Supporting Table 2**).

Subsequently, a data-mining search was performed to identify whether the SANGER-identified sequence 81 was present in the library pool. For this reason, we mined the RV top-enriched dataset for a specific sequence tag that identified sequence 81 within the chimeric library (5'-GTTGAGTAGAG-TGTGAGCTTGTCACACATGTGTTGACGACTCGCTGTTGTTACAGTGTTACACA TAGCCGT-3'), but no hit was identified.

	Index	Sequence count	Different sequences	Sequence variability
Starting pool	CTTGTA	223781	218097	97%
Cycle 2	ACAGTG	1076127	158552	14%
Cycle 3	GCCAAT	515000	53208	10%

Table 6 | NGS indices and sequence representation for the µIVC screening on the chimeric library.

5.6.3 Specs NGS analysis repetition of µIVC screening

The repetition of the NGS analysis was conducted using the same sequencer and chemistry of the previous analysis but with the NextSeq 150 cycles sequencing kit to sequence the whole library from both FW- and RV-primer sides. Also, all selection cycles were sequenced this time. The raw sequencing results were mined to identify only the full-length library sequences (AACAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTG-TGAGCTNNNNNCACANGTGTTGACGACTCGCTGNNNNN) and to discard the enriched truncated sequences, which cause problems in the previous NGS runs. The top-50 enriched sequences of cycle 3 gate 2 (12.84%-0.11%) were aligned to the library pool (**Supporting Table 3, Supporting Table 4**).

	Index	Sequence count	Different sequences	Sequence variability
Starting pool	ACAGTG	4689186	2442032	59%
Cycle 1	GCCAAT	1196850	143796	12%
Cycle 2	CAGATC	679960	16972	2,5%
Cycle 3 gate 1	ACTTGA	398498	13709	3,4%
Cycle 3 gate 2	TAGCTT	659587	8300	1,26%
Cycle 3 gate 3	GATCAG	316688	3974	1,25%

Table 7 | NGS sequencing repetition indices and sequence representation for the µIVC screening on the chimeric library.

5.7 Sanger sequencing

After µIVC screening, the enriched library was cloned into vectors and transformed into competent cells with the TOPO-TA Cloning kit according to the manufacturer's recommendation. Before cloning, the PCR sample from screening cycle 3 was re-amplified with GoTaq DNA polymerase with a final 20 minutes 72 °C step to produce 3' overhang deoxyadenosine crucial for the cloning strategy. After plasmid and clones work-up, 30 single DNA clones were sequenced by Sanger seq. (Microsynth, Balgach, Switzerland) with the M13 reverse primer.

5.8 Expression of the sensor in mammals

DNA templates of the scaffolded Sensor and Control were inserted in the vector pAV-U6+27-Tornado-F30-Broccoli-empty, published and kindly provided by Litke *et al.*[135] The plasmid map is presented in **Supporting Figure 7**. In essence, the vector allows the transcription of an RNA insert under the control of a U6+27 promoter with a U6 terminator; the pAV vector contains an SV40 origin and ampicillin selection cassette. To be inserted, the scaffolded Sensor and Control

were initially PCR amplified with two primers (FW: 5'-CGGCCGCTTGCCATGTGTATGTGGGTA -3'; RV: 5'- GGCCACTTGCCATG-AATGATCCCGAAG -3'). Designed with ApE software, these primers were used to obtain a 25-nt homologous flanking regions between the vector and the inserts. The inserts were cloned into the vector following AQUA cloning protocol.[191] 12 fmol of Backbone were resuspended in a final volume of 5 µL H_2O with 120 fmol of Sensor or Control inserts (1:10 molar ratio) and incubated 1 hr at 25°C. The solution was used to transform HST08 stellar *E. coli* cells (Takara bio no. 636763). After a short incubation in ice, a mixture of chemically competent bacteria and DNA were placed at 42°C for 45 seconds (heat shock) and placed back in ice. SOC medium was added, and the transformed cells were incubated at 37°C for 1 hr with agitation and plated on agar plates with 100 µg/mL ampicillin at 37°C overnight. Single colonies were inoculated in LB medium with 100 µg/mL ampicillin overnight, and plasmid DNA was extracted using the NucleoBond Xtra Midi kit. Constructs were sequence verified by Sanger sequencing (Microsynth).

5.9 Cell culture and transfection

HEK293T (ATCC CRL-3216), Hela (ATCC CCL-2), and HepG2 (ATCC HB-8065) cells were maintained in DMEM, 10% FBS, 100 U/mL penicillin, 100 mg/mL streptomycin solution under standard tissue culture condition. All cells were split using TrypLE Express according to the manufacturer's instructions. Cell lines were plated in a 24-well TC plate and transfected using FuGENE using Opti-MEM I Reduced Serum Media to resuspend the vector. To promote HEK293T adhesion, the plate was treated with Poly-L-Lysine Solution for 1 hr and washed with D-PBS. For HEK293T cells, 5×10^4 cells were seeded per well and transfected with a 25 µL Opti-MEM solution containing 500 ng and 1.5 µL transfection reagent per well. For Hela cells, 4×10^4 cells were seeded per well and transfected with a 25 µL Opti-MEM solution containing 500 ng and 1.5 µL transfection reagent per well. Finally, for HepG2 cells, 10^5 cells were seeded per well and transfected with a 25 µL Opti-

MEM solution containing 500 ng and 2.3 µL transfection reagent per well. All cells were transfected 24 hr after seeding.

5.10 Flow cytometry comparing constructs in different cell lines

Two days after transfection, HEK293T, Hela, and HepG2 cells were detached with TrypLE, diluted with DMEM 10% FBS, and centrifuged at 200 rcf for 5 minutes. Cellular pellet was resuspended in DMEM 10% FBS containing 40 µM DFHBI-1T and kept on ice till analysis. Cells were analyzed using FACS Canto II (BD Bioscience). The population of cells was gated to avoid cell doublets detected by the forward and side scattering; the gating strategy is shown in **Supporting Figure 8**. Plots were generated using FlowJo software v9.6.3.

5.11 Oligomycin A treatment

Analogous to the constructs comparison, HEK293T cells were seeded on a 24-well plate in DMEM 10% FBS and transfected the next day. Two days after transfection, the cells medium was switched to a formulated DMEM (1 g/L Glucose, 10% FBS) containing 2 mM Glutamine, added shortly before the assay. The formulated medium was prepared according to the manufacturer's instructions for powdered DMEM. For positive cells, the cell medium was supplemented with 10 µM oligomycin A for 30 min. Cells were detached with TrypLE, diluted with the corresponding DMEM formulation, and centrifuged at 200 *g* for 5 min at 4°C. Cellular pellet was resuspended in corresponding DMEM formulation containing 40 µM DFHBI-1T and kept on ice until analysis. Cells were gated and normalized analogous to constructs comparison.

The oligomycin assay was repeated in a double-blind setting to ensure non-biased sample treatment. Samples were prepared for transfection and handed out to a second investigator for relabeling. Samples were transfected, and the oligomycin

assay was performed according to protocol. After flow cytometry measurement, the sample blinding was unveiled.

5.12 LC-MS Metabolite quantification

The metabolite extraction followed the protocol described by Park et al., 2019[157]. Culture dishes of 150 cm2 (734–2322, VWR) were treated with poly-lysine solution (20 mL/plate) for 1hr and then washed twice with DPBS and kept under the hood to dry up. 6 x 10^6 HEK293T cells were seeded in each plate and cultivated for 48 hr in formulated DMEM (1 g\L Glucose, 2 mM Glutamine, 10% FBS).

Three dishes were seeded for each condition: oligomycin (treated), DMSO (untreated), pre-spiked control, post-spiked control. 1 dish was used for cell counting, whereas 2 dishes were subjected to metabolite extraction. After 48-hour growth, cells were treated with a fresh formulated medium containing oligomycin (10 µM) or DMSO vehicle (untreated) for 30 minutes. Pre-spiked and post-spiked controls were also treated with DMSO vehicle. Subsequently, cell metabolism was quenched by placing the cell dishes on a -20°C metal block and the respective metabolite extraction protocol. Mammalian cell metabolism was quenched, and metabolites were extracted by quickly aspirating media from culture dishes and adding cold (–20 °C) 80:20 methanol/water. The plates were moved into a –20 °C freezer for 20 min. Then, cells were scraped off the culture dish surface. Note that, in all cases, quenching was performed without any washing steps that can perturb the metabolism.

The dishes were rinsed with an additional 15 mL of ice-cold ddH2O, and the two suspensions were collected together and quenched in LN2. Before LN2 quenching, the pre-spiked control added a metabolite mixture (1mM G6P, 1mM FMP, and 1 mM FBP) of 400 uL. The samples were cleared from cellular debris by centrifugation (20 mins at 4 °C, 4500× g). At this point, also the post-spiked control solution was spiked with a metabolite mixture (1mM G6P, 1mM FMP, and 1 mM

FBP) of 400 uL. The 30 mL solution was frozen at -80°, and the frozen metabolite extracts were furthermore concentrated by lyophilization and stored at −80 °C.

The metabolite extract analysis was performed by the team of prof. dr. Mayer (University of Tübingen). The metabolite extracts were analyzed using ESI-TOF (micrO-TOF II, Bruker) mass spectrometer connected to an UltiMate 3000 high-performance liquid chromatography. For all metabolites beside FBP, 5 µl of standards or samples were injected onto HILICpak VT-50 2D column (Shodex), and MS analysis was performed in negative-ion mode over the mass range from 200 to 1,000 m/z. Instead, FBP was run in a SEQuant ZIC-pHILIC column (Merck), and MS analysis was performed in high-pH gradient mode over the mass range from 200 to 1,000 m/z. Extracted ion chromatogram (EIC) spectra were plotted in Data Analysis (Bruker).

5.13 FBP colorimetric assay

The metabolite extraction used to quantify FBP via colorimetric assay was performed head-to-head with the oligomycin-assay, which was analyzed by flow cytometry (**Figure 32**, protocol as in 5.11) using the same passage of HEK293T cells.

HEK293T cells were handled similarly to the metabolite quantification by LC-MS to obtain the metabolite extracts, with some differences. Culture dishes of 21 cm^2 (150462, Thermo Fisher) were treated with poly-lysine solution (5 mL/plate) for 1 hr, washed twice with DPBS, and kept under the hood to dry up. 1 x 10^6 HEK293T cells were seeded in each plate and cultivated for 48 hr in formulated DMEM (1 g\L Glucose, 2 mM Glutamine, 10% FBS). 5 dishes were seeded for each condition: oligomycin (treated), DMSO (untreated). 2 dishes were used for cell counting, whereas 3 dishes were subjected to metabolite extraction. After 48-hour growth, cells were treated with a fresh formulated medium containing oligomycin (10 µM) or DMSO vehicle (untreated) for 30 minutes. The medium was then quickly removed,

and cells were scraped and collected in 1 mL FBP assay buffer, provided by the supplier.

The cell suspension was homogenized in a bead homogenizer (Omni International) following the protocol provided by Mathon *et al*.[192] Briefly, one metal bead was added for each sample, and the cell suspension was crushed for 2 min at 30 Hz in -20°C cooled blocks. The cell debris was precipitated by centrifugation (16800g x 10 min). Finally, the supernatant was run through a 3K centrifugation filter to harvest the metabolites; this solution was used for the colorimetric assay.

For each sample, 35 µL purified metabolite solution was then incubated with 2 µL FBPase (enzyme provided by the kit and resuspended in ddH_2O+ 30% glycerol) 5 hr at 37°C to convert all FBP to FMP. The spiked control samples were supplied with FBP at increasing final concentration (0.01, 0.1, and 1 mM) before this step to confirm the enzyme activity. Instead, to measure the background level of FMP. Also, 35 µL of each sample was incubated at 37°C for 5 hr only with 2 µL ddH_2O+ 30% glycerol to ensure that the enzyme-treated and untreated samples were handled in the same way.

Afterward, a converter, developer, and probe (CDP) solution was prepared according to the manufacturer's protocol and kept at 4°C before use. 8 µL of CDP solution were added to each 35 µL sample. 20 µL of each sample were plated in a 384 absorbance plate (EP0030128508, Eppendorf) and the end-point absorption at 450nm was obtained. To isolate the signal contribution of the FBP converted to FMP, the background absorbance (samples treated without FBPase) was subtracted to the FBPase-treated samples, thus obtaining the ΔAbs values, which were plotted in **Figure 32**. Absorbance was measured using the EnSpire Multimode Plate Reader (PerkinElmer).

6 MATERIALS

6.1 Oligonucleotides

All oligonucleotides were purchased as HPLC-purified single-stranded DNA in a 0.4 micromole scale from Ella Biotech (Planegg, Germany). All libraries' random regions were synthesized with 1:1:1:1 ratio of each nucleotide.

NAME	SEQUENCE
Stem-loop primed library	5'-GGACGTTATACACTGTATG-N26-GTAGTTCGCTAC-N26-ACATAACTTGGTAAATGTGCTT-3'
S0 library	5'-GGACGTTATACACTGTATGTGGTGAAGGACGGGTNNNGTAGTTCGCTAC-NNNNNNGTAGAGTGTGAGCTCCACATAACTTGGTAAATGTGCTT-3'
S1 library	5'-GGACGTTATACACTGTATGTGGTGAAGGACGGGTCCAG-N20-GTTCGC-N20-CTGTTGAGTAGAGTGTGAGCTCCACATAACTTGGTAAATGTGCTT-3'
S2 library	5'-GGACGTTATACACTGTATGTGGTGAAGGACGGGTC-N20-GTTCGC-N20-TTGAGTAGAGTGTGAGCTCCACATAACTTGGTAAATGTGCTT-3'
S3 library	5'-GATGTAACTGAATGAAATGGTGAAGGACGGG-N40-GTAGAGTGTGAGCTCCGTAACTAGTTACATCCATCGCACGAGACCT-3'
S4 library	5'-GGTAGAGCACAATAAGTCAA-N20-GGTGAAGGACGGGTCCAGTAGTTC-

	GCTACTGTTGAGTAGAGTGTGAGCTCC-N20-TTGACACAGTGTTACACAT-AGCCGT-3’
S5 library	5'-GGCAGCCTACTTGTTGAGTAGAGTGTGAGC-N20-GAAGGACGGGTCCAGT-AGGCTGCTTCTACCTGATGACGAA -3'
S6 library	5’-GGTAGAGCACAATAAGTCAA-N18-GGANGGGTCCAGTAGTTCGCTACT-GTTGAGNAGAGTG-N23-TTGACACAGTGTTACACATAGCCGT-3’
Chimeric library	5’-GGTAGAGCACAATAAGTCAACAAGGACGGGTCCAGTAGTTCGCTACTGTT-GAGTAGAGTGTGAG-CT-N5-CACANGTGTTGACGACTCGCTG-N5-ACAGTGTTACACATAGCCGT-3’
D1 aptamer	5’-GGACGTTATACACTGTATGTGGTGAAGGACGGGTCCAGTAGTTCGCTACT-GTTGAGTAGAGTGTGAGCTCCACATAACTTGGTAAATGTGCTT -3’
Seq 81	5’-GGTAGAGCACAATAAGTCAAGAAGGACGGGTCCAGTAGTTCGCTACTGTT-GAGTAGAGTGTGAGCTTGTCACACATGTGTTGACGACTCGCTGTTGTTACAG TGTTACACATAGCCGT-3’
1 point mutant,1PM	5’-GGTAGAGCACAATAAGTCAAGAAGGACGGGTCCAGTAGTTCGCTACTGTT-GAGTAGAGTGTGAGCTTGTCATACATGTGTTGACGACTCGCTGTTGTTACAGT GTTACACATAGCCGT-3’
2 point mutant, 2PM	5’-GGTAGAGCACAATAAGTCAAGAAGGACGGGTCCAGTAGTTCGCTACTGTT-GAGTAGAGTGTGAGCTTGTCATATATGTGTTGACGACTCGCTGTTGTTACAGT GTTACACATAGCCGT-3’
4 point mutant, 4PM, Control	5’-GGTAGAGCACAATAAGTCAAGAAGGACGGGTCCAGTAGTTCGCTACTGTT-GAGTAGAGTGTGAGCTTGTCATGTTTGTGTTGACGACTCGCTGTTGTTACAGT GTTACACATAGCCGT-3’
Scaffolded Sensor	5’-GGTTGCCATGTGTATGTGGGTAAGTCAAGAAGGACGGGTCCAGTAGTTCG-CTACTGTTGAGTAGAGTGTGAGCTTGTCACACATGTGTTGACGACTCGCTGTT GTTACAGTGTTACACATAGCCCACATACTCTGATGATCCTTCGGGATCATTCA TGGCAA-3’
Scaffolded 4PM control	5’-GGTTGCCATGTGTATGTGGGTAGAGCACAATAAGTCAAGAAGGACGGGTC-CAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCTTGTCATGTTTGTGTTGACG ACTCGCTGTTGTTACAGTGTTACACATAGCCCACATACTCTGATGATCCTTCG GGATCATTCATGGCAA-3’
FW primer for	5’-TAATACGACTCACTATAGGACGTTATACACTGTATGT-3’

S0, S1, S2 libraries and D1 aptamer	
RV primer for S0, S1, S2 libraries and D1 aptamer	5’-AAGCACATTTACCAAGTTA-3’
FW primer for S4. S5, S6 libraries and seq 81 plus 1-4PMs	5’-TAATACGACTCACTATAGGTAGAGCACAATAAGTCAA-3’
RV primer for S4. S5, S6 libraries and seq 81 plus 1-4PMs	5’-ACGGCTATGTGTAACACTGT-3’
FW primer S3 library	5'-CCAAGTAATACGACTCACTATAGATGTAACTGAATGAAAT-3'
RV primer S3 library	5'-AGGTCTCGTGCGATGGATGT-3'
FW primer S5 library	5'-CCAAGTAATACGACTCACTATAGGCAGCCTACTTGTT-3'
RV primer S5 library	5'-TTCGTCATCAGGTAGAAGCA-3'
FW primer for scaffolded sensor and control	5’-TAATACGACTCACTATAGGTTGCCATGTGTATGT-3’
RV primer for scaffolded sensor and control	5’- TTGCCATGAATGATC -3’

6.2 Molecular biology reagents

NAME	SOURCE	IDENTIFIER
1,4-dithiothreitol (DTT)	AppliChem	A1101
1H,1H,2H,2H-Perfluoro-1-octanol	Merck	370533
Adenosine 5′-monophosphate disodium salt	Sigma	01930
Agar Bacteriology grade	AppliChem	A0949
Agarose LE	Genaxxon	M3044
Ampicillin sodium salt (Amp)	Carl Roth	HP62.1
Boric acid	AppliChem	A2940
Chloroform	Thermo Fisher Scientific	C298-500
Cytidine 5′-monophosphate disodium salt	Sigma	C1006
DFHBI	Merck	SML1627
D-Fructose 1,6-bisphosphate trisodium salt hydrate (FBP)	Merck	F6803
D-Fructose 6-phosphate disodium salt hydrate (FMP)	Merck	F3627
D-Glucosamine 6-phosphate sodium salt (GluN6P)	Merck	G4878
D-Glucose (Glu)	Merck	NIST917C
D-Glucose 6-phosphate disodium salt hydrate (Glu6P)	Merck	G7250
Dimethyl sulfoxide	Merck	472301
DNase I (RNase-free)	NEB	M0303
dNTP Bundle 100 mM (dATP, dCTP,dGTP, dTTP)	Jena Bioscience	NU-1005
Enhanced Avian Reverse Transcriptase	Merck	A4464

[eAMV™ RT]		
Ethanol absolute	Merck	32205
Ethidium bromide solution (1.00%)	Carl Roth	2218
FACS Clean solution	Becton Dickinson	340345
FACS Flow solution	Becton Dickinson	342003
FACS Shutdown solution	Becton Dickinson	334224
Gel Loading Dye, Purple (6x), no SDS	NEB	B7025S
GeneRuler 100 bp DNA Ladder	Thermo Fisher Scientific	SM0241
Glycerol (86%)	Carl Roth	4043
GoTaq® DNA Polymerase	Promega, Madison	M3008
Guanosine 5′-monophosphate disodium salt hydrate	Sigma	G8377
HEPES	Carl Roth	HN78.1
Hydrochloric acid (HCl)	Carl Roth	231-595-7
Lennox Broth (LB)	Carl Roth	X964
Methylamine solution 33 wt. %	Merck	741851
Molecular Probes™ Dextran Texas Red™, 70,000 MW	Thermo Fisher Scientific	D1864
NHS-Activated Sepharose 4 Fast Flow, 25 mL	GE Healthcare	17090601
NTP Bundle 100 mM (ATP, CTP, GTP, UTP)	Jena Bioscience	NU-1014
Potassium acetate (KAc)	Merck	3595449
RNasin® Ribonuclease Inhibitors	Promega	N2511
Phenol (Roti-Phenol)	Carl Roth	0038.3
Pluronic® F-68	Merck	P1300
Pyrophosphatase, inorganic (PPase)	Merck	10108987001
SOC Medium	Merck	S1797

Sodium acetate (NaOAc)	Merck	W302406
Sodium Chloride (NaCl)	Thermo Fisher Scientific	J21618
Spermidine	Merck	S0266
T7 polymerase	homemade	
Thymidine 5′-monophosphate disodium salt hydrate	Sigma	T7004
Tris	Carl Roth	5429.1
Urea	Carl Roth	2317.2
Uridine 5′-monophosphate disodium salt	Sigma	U6375
Yeast tRNA	Thermo Fisher Scientific	AM7119

6.3 Cell culture reagents

NAME	SOURCE	IDENTIFIER
0.05% Trypsin-EDTA (1X)	Thermo Fisher Scientific	25300054
Cell culture plates and flasks	Sarstedt	n/a
DMEM (1X) + GlutaMAX-I	Thermo Fisher Scientific	61965026
Dulbecco′s Modified Eagle′s Medium Powder	Merck	D5030
DFHBI-1T	Bio-Techne, Minneapolis	5610
DPBS (1X)	Thermo Fisher Scientific	14190144
Fetal bovine serum (FBS)	Merck	F9665
FuGENE HD Transfection Reagent	Promega	E2311
L-Glutamine	Sigma	G6392
Oligomycin A	Tocris	4110
Opti-MEM I (1X)	Thermo Fisher Scientific	31985062
Penicillin/Streptomycin (100X)	PAA	P11-010

Poly-L-lysine solution (0.01% in H2O)	Merck	A-005-C
Sodium Bicarbonate	Merck	6014

6.4 Critical commercial kits

NAME	SOURCE	IDENTIFIER
Fructose-1,6-Bisphosphate Assay Kit	Biovision	K590
MicroSpin™ G-25 Columns	GE Healthcare	45-001-397
NucleoSpin Gel and PCR Clean-up	Macherey-Nagel	740609
NucleoBond Xtra Midi	Macherey-Nagel	740588
TOPO TA cloning kit with One Shot TOP10 chemically competent E. coli	Thermo Fisher Scientific	K450001
TruSeq DNA PCR-Free Sample Preparation Kit LT	Illumina	20015963

7 SUPPLEMENTS

SUPPORTING TABLES

	Enriched sequences aligned to the initial pool	RNA length	Frequency [%] c0	c2	c3G2
Pool	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCT**NNNNN**CACA**N**GTGTTGACGACTCGCTG**NNNNN**	118			
	CAAGGACGGGTCCAGTAG-T----AACTGTTGAGTAGAGTGTG----------CCTCAGGGTTTCGCATTATTGTTG	102	0.00	0.89	3.28
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG----------TGTTGACGACTCGCTGAATGT	105	0.00	0.18	3.05
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG----------TGTTGACGACTCGCTGAAGTA	105	0.00	0.56	2.15
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG----------TGTTGACGACTCGCTGCGGGA	105	0.00	0.42	1.90
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG----------TGTTGACGACTCGCTGCAAGA	105	0.00	0.50	1.34
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTG----------GTTTTGACGACTCGCTGTAGAG	104	0.00	0.30	1.19
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTG----------GGTGTTGACGACTCGCTGCTCTT	103	0.00	0.19	1.08
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGT----------GGGGCGACTTGGATCGTGGCAATAGAC	108	0.00	0.30	0.96
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGTGT----------TGACGACTCGCTGCATTT	105	0.00	0.18	0.95
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGTGT----------TGACGACTCGCTGTACAA	105	0.00	0.36	0.87
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTG----------TGTGTTGACGACTCGCTGCTCGC	105	0.00	0.25	0.77
X	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGAAGCACATGTGTTGACGACTCGCTGATCCG	118	0.00	0.20	0.75
85	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGCAACACAAGTGTTGACGACTCGCTGCGGGA	118	0.00	0.35	0.74
	CAAGGACGGGTCCAGTAG-TT----ACTGTTGAGTAGAGTGT----------GCCTCAGGGTTTCGCATTATTGTTG	102	0.00	2.09	0.70
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGACGCACAGGTGTTGACGACTCGCTGAAACT	118	0.00	0.14	0.67
86	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGTCGCACAAGTGTTGACGACTCGCTGTTACG	118	0.00	0.18	0.54
	GAAGGACGGGTCCAGTAGGTTCGCTACTGTTGAGTAGAGTGTGAGCTGACTTCACATGTGTTGACGACTCGCTGGGCAG	119	0.00	0.25	0.49
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGGTTGACACACACGTGTTGACGACTCGCTGTATGG	118	0.00	0.22	0.46
	GAAGGACGGGTCCAGTAGGTTCGCTACTGTTGAGTAGAGTGTGAGCTGACTTCACAGGTGTTGACGACTCGCTGGTCAG	119	0.00	0.60	0.42
	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTTGACCACATGTGTTGACGACTCGCTGGTATA	118	0.00	0.29	0.37
	GAAGGACGGGTCCAGTA-TTTTCGCTACTGTTGAGTAGAGTGTGAGCTTGACTCACATGTGTTGACGACTCGCTGGGTTG	118	0.00	0.26	0.27
	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGTACCACAGGTGTTGACGACTCGCTGTAGTC	118	0.00	0.58	0.24
	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGTATCACAGGTGTTGACGACTCGCTGTTTAG	118	0.00	0.60	0.23
	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGCTGCACATGTGTTGACGACTCGCTGTACAT	118	0.00	0.18	0.22

Supporting Table 1 | Top enriched sequences from cycle 3.G2 of µIVC screening identified by NGS analysis.
The conserved motifs on the sequences are highlighted (red), and their RNA length is reported together with the enrichment frequency per cycle. Shortlisted sequences are highlighted (grey).

	Sanger sequenced clones aligned to the initial pool	RNA length
Pool	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTNNNNNCACANGTGTTGACGACTCGCTGNNNNN	118
53	-AAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTGACTTCACAGGTGTTGACGACTCGCTGCCCAT	117
77	-AAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTGACTTCACAGGTGTTGACGACTCGCTGCCCAT	117
52	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTGACTACACAAGTGTTGACGACTCGCTGCTGGT	118
55	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGATTAGAGTGTGAGCTTGGTTAACAGGTGTTGACGACTCGCTGCCCGG	118
74	AAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGGTTAACAGGTGTTGACGACTCGCTACCCGG	118
81	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGTCACACATGTGTTGACGACTCGCTGTTGTT	118
86	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGTCGCACAAGTGTTGACGACTCGCTGTTACG	118
60	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGGTTGAGTCACATATGTTGACGACTCGCTGAGGCT	118
62	CAAGGACGGATCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGACCCACAGGTATTGACGACTCGCTGGAATA	118
82	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGGTTGATACACAGGTGTTGACGACTGGCTGCCGTT	118
76	CAAGGACGGGTCCAGTAA-TTCGCTACTGTTGAGTAGAGTGTGAGCTTGAAGCACATGTGTTGACGACTCGCTGGGGGT	118
85	GAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTGGAGTGTGAGCTTGCAACACAAGTGTTGACGACTCGCTGCGGGA	118
73	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTATTTACACACGTGTTGACGACTTGCTGATGTG	118
88	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAGCTTCTAGCACATGTGTTGACGACTCGCTGTAACT	118
75	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGATTGTGGGCTCTCTGCACACGTGTTGACGACTCNNTGATGTA	118
63	CAAGGACGGATCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG--------------GTTGACGACTCGCTGATTGC	104
64	CAAGGACGGATCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG--------------GTTGACGACTCGCTGATTGC	104
84	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG-------------TGTCGACGACTCGCTGTACGA	105
65	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG-------------TGTTGACGACTCGCTGCATTT	105
90	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGAG-------------TGTTGACGACTCGATGCAAGA	105
54	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTG--------------GTTTTGACGACTCGCTGTAAAG	104
66	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTG-------------CCTGGTGCGAATGTGTGTGTTGA	105
79	CAAGGACGGGTCCAGTAG-TTCGCTACTGTTGAGTAGAGTGTGA--------------TGTTGACGACTCGATGGGGTA	104
83	CAAGGACGGGTCCAGTAG-TTCGCTANTGTTGAGTAGAGTGTGA--------------TGTTGACGACTCGCTGCAGAT	104
57	CAAGGACGGGTCCAGTAG-TTCGCTACGGTTGAGTAGAGTGTGA--------------GGTTGACGACTCGCTGAAGCC	104
70	CAAGGACGGGTCCAGTAT-TTCGCTACTGTTGAGTAGAGTGTGA--------------TGTTGACGACTCGCTGCATTG	104
69	CAAGGACGGGTCCAGTTG-TTCGATACTGTTGAGCAGAGCGTGA--------------GGTTGACGATTCGCTGTAGGA	104

Supporting Table 2 | Clones identified by Sanger sequencing on cycle 3.G2 of µIVC screening.

The conserved motifs on the sequences are highlighted (red), and their RNA length is reported. Shortlisted sequences are highlighted (grey).

		CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCTNNNNNCACANGTGTTGACGACTCGCTGNNNNN
	A1	AAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGACG**CACA**G**GTGTTGACGACTCGCTG**AAACT**
53	A2	AAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**GACTT**CACA**G**GTGTTGACGACTCGCTG**CCCAT**A
X	A3	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGAAG**CACA**T**GTGTTGACGACTCGCTG**ATCCG**
85	A4	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGCAA**CACA**A**GTGTTGACGACTCGCTG**CGGGA**
	A5	AAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTGAC**CACA**T**GTGTTGACGACTCGCTG**GTATA**
	A6	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**GACTA**CACA**A**GTGTTGACGACTCGCTG**CCGGT**
	A7	AAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAG-CT**GACTT**CACA**T**GTGTTGACGACTCGCTG**GGCAG**A
86	A8	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTCG**CACA**A**GTGTTGACGACTCGCTG**TTACG**
	A9	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGGAA**CACA**T**GTGTTGACGACTCGCTG**AATCC**
	A10	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTAT**CACA**G**GTGTTGACGACTCGCTG**TTTAG**
	A11	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGGGG**CACA**T**GTGTTGACGACTCGCTG**TATTT**
	A12	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGCTG**CACA**T**GTGTTGACGACTCGCTG**TACAT**
	A13	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGAAA**CACA**G**GTGTTGACGACTCGCTG**CATCT**
	A14	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTCC**CACA**G**GTGTTGACGACTCGCTG**TACGA**
	A15	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTTGT**CACA**A**GTGTTGACGACTCGCTG**TGTAG**
	A16	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGGCC**CACA**A**GTGTTGACGACTCGCTG**TGCGG**
	A17	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTAG**CACA**A**GTGTTGACGACTCGCTG**TTACG**
	A18	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTAC**CACA**G**GTGTTGACGACTCGCTG**TAGTC**
	A19	AAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCTG**ACTTC**ACAG**G**TGTTGACGACTCGCTGC**ACATA**
	A20	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTTGT**CACA**A**GTGTTGACGACTCGCTG**TGTCC**
	A21	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTCA**CACA**T**GTGTTGACGACTCGCTG**ATGTT**
	A22	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGCAA**CACA**A**GTGTTGACGACTCGCTG**AGGGA**
	A23	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTAC**CACA**G**GTGTTGACGACTCGCTG**CATAT**
	A24	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGACC**CACA**C**GTGTTGACGACTCGCTG**ATTAC**
	A25	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGAAG**CACA**T**GTGTTGACGACTCGCTG**GGGGG**
	A26	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGGTT**CACA**G**GTGTTGACGACTCGCTG**TTGAA**
	A27	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGCTG**CACA**T**GTGTTGACGACTCGCTG**AGAAT**
	A28	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGAAC**CACA**G**GTGTTGACGACTCGCTG**CATCT**
	A29	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TAGAC**CACA**T**GTGTTGACGACTCGCTG**ATAGT**
	A30	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTTGT**CACA**A**GTGTTGACGACTCGCTG**TGTCG**
	A31	AAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCTG**ACTTC**ACAT**G**TGTTGACGACTCGCTGG**GCATA**
	A32	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGCTT**CACA**T**GTGTTGACGACTCGCTG**AGGGT**
	A33	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTGAG**CACA**A**GTGTTGACGACTCGCTG**TTTTA**
	A34	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGACG**CACA**G**GTGTTGACGACTCGCTG**AAACT**
	A35	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTCGA**CACA**T**GTGTTGACGACTCGCTG**GTAAA**
	A36	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGATT**CACA**T**GTGTTGACGACTCGCTG**CCTGT**
	A37	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTGTG**CACA**A**GTGTTGACGACTCGCTG**TAAGC**
	A38	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**GACTT**CACA**G**GTGTTGACGACTCGCTG**CCCAT**
	A39	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**GGCTT**CACA**G**GTGTTGACGACTCGCTG**TAAAC**
	A40	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**GTATT**CACA**A**GTGTTGACGACTCGCTG**CTGGA**
	A41	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**AAACT**CACA**A**GTGTTGACGACTCGCTG**GTGTT**
	A42	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGATT**CACA**T**GTGTTGACGACTCGCTG**GTGAT**
	A43	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**GATCT**CACA**G**GTGTTGACGACTCGCTG**AATGC**
	A44	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTGAC**CACA**T**GTGTTGACGACTCGCTG**GTATA**
	A45	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGAAG**CACA**T**GTGTTGACGACTCGCTG**ATCCG**
	A46	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**AGAAT**CACA**T**GTGTTGACGACTCGCTG**GACAG**
	A47	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTAA**CACA**A**GTGTTGACGACTCGCTG**CGCAC**
81	A48	GAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TGTCA**CACA**T**GTGTTGACGACTCGCTG**TTGTT**
	A49	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**TTGGC**CACA**G**GTGTTGACGACTCGCTG**CCTGG**
	A50	CAAGGACGGGTCCAGTAGTTCGCTACTGTTGAGTAGAGTGTGAGCT**CGACT**CACA**T**GTGTTGACGACTCGCTG**CCCGG**

Supporting Table 3 | Top-50 enriched sequences obtained by NGS of the µIVC screening. The top-50 sequences enriched in cycle 3.G2 of µIVC screening were aligned to the chimeric library (red). The common sequences with the previous NGS/SANGER analysis are marked (grey boxes); the sequence 81 position (A48) is highlighted (yellow).

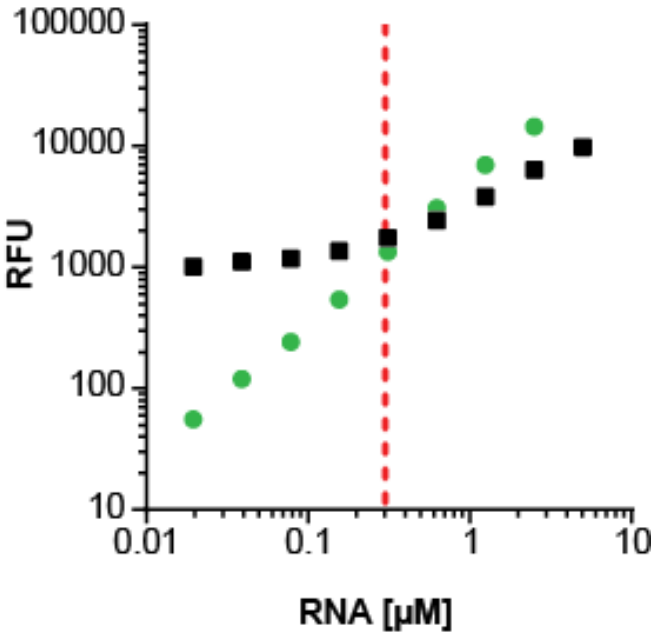
	c0	c1	c2	c3G1	c3G2	c3G3
A1	0.00	0.07	3.86	0.20	12.84	0.33
A2	0.00	0.06	3.39	0.61	11.03	0.00
A3	0.00	0.04	1.70	0.27	8.89	0.68
A4	0.00	0.06	2.71	0.06	8.00	0.35
A5	0.00	0.08	2.55	0.15	6.65	0.20
A6	0.00	0.05	2.88	0.06	5.19	0.64
A7	0.00	0.07	1.93	0.02	4.89	0.00
A8	0.00	0.02	1.28	0.22	4.47	0.02
A9	0.00	0.01	1.03	0.01	3.92	0.01
A10	0.00	0.08	5.84	0.05	2.85	0.02
A11	0.00	0.02	0.74	0.00	2.09	0.00
A12	0.00	0.07	4.53	0.07	1.88	0.03
A13	0.00	0.00	0.45	0.01	1.23	0.01
A14	0.00	0.06	2.41	0.06	0.75	0.08
A15	0.00	0.00	0.31	0.02	0.67	0.00
A16	0.00	0.03	1.96	0.04	0.66	0.47
A17	0.00	0.00	0.28	0.01	0.56	0.00
A18	0.00	0.04	4.00	0.26	0.55	0.71
A19	0.00	0.00	0.21	0.01	0.53	0.00
A20	0.00	0.00	0.30	0.00	0.48	0.00
A21	0.00	0.00	0.29	0.00	0.46	0.00
A22	0.00	0.00	0.28	0.00	0.44	0.00
A23	0.00	0.03	2.25	0.04	0.42	0.02
A24	0.00	0.03	1.99	0.36	0.40	0.02
A25	0.00	0.01	0.87	0.01	0.39	0.03

	c0	c1	c2	c3G1	c3G2	c3G3
A26	0.00	0.03	1.31	0.03	0.38	0.01
A27	0.00	0.00	0.48	0.00	0.37	0.00
A28	0.00	0.02	0.49	0.02	0.32	0.00
A29	0.00	0.04	3.25	0.06	0.29	0.14
A30	0.00	0.09	0.54	0.21	0.28	0.00
A31	0.00	0.00	0.19	0.00	0.25	0.00
A32	0.00	0.06	3.49	0.06	0.24	0.28
A33	0.00	0.03	1.41	0.11	0.24	0.33
A34	0.00	0.00	0.08	0.02	0.23	0.01
A35	0.00	0.02	1.12	0.01	0.20	0.01
A36	0.00	0.03	1.35	0.10	0.18	16.39
A37	0.00	0.00	0.13	0.82	0.18	0.00
A38	0.00	0.00	0.06	0.03	0.18	0.00
A39	0.00	0.01	0.57	0.05	0.17	0.01
A40	0.00	0.00	0.13	0.36	0.17	0.00
A41	0.00	0.00	0.10	0.66	0.16	0.00
A42	0.00	0.01	0.72	0.01	0.16	10.04
A43	0.00	0.00	0.27	1.39	0.16	0.00
A44	0.00	0.00	0.05	0.02	0.15	0.05
A45	0.00	0.00	0.03	0.01	0.15	0.01
A46	0.00	0.09	0.32	2.20	0.13	0.00
A47	0.00	0.00	0.18	0.00	0.13	0.00
A48	0.00	0.01	0.78	0.01	0.13	0.27
A49	0.00	0.00	0.08	0.39	0.12	0.00
A50	0.00	0.06	0.12	0.40	0.11	0.00

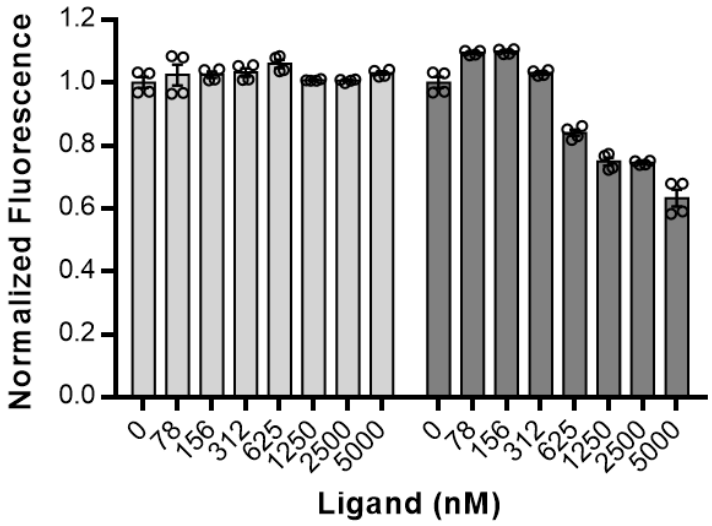
Supporting Table 4 | Enrichment frequency of the top-50 enriched sequences obtained by NGS of the µIVC screening.

The table reports the percentile frequency per screening cycle of the top-50 sequences listed in **Supporting Table 3**. The sequences are organized according to their enrichment in cycle 3.G2. The position of sequence 81 (A48) is highlighted (yellow).

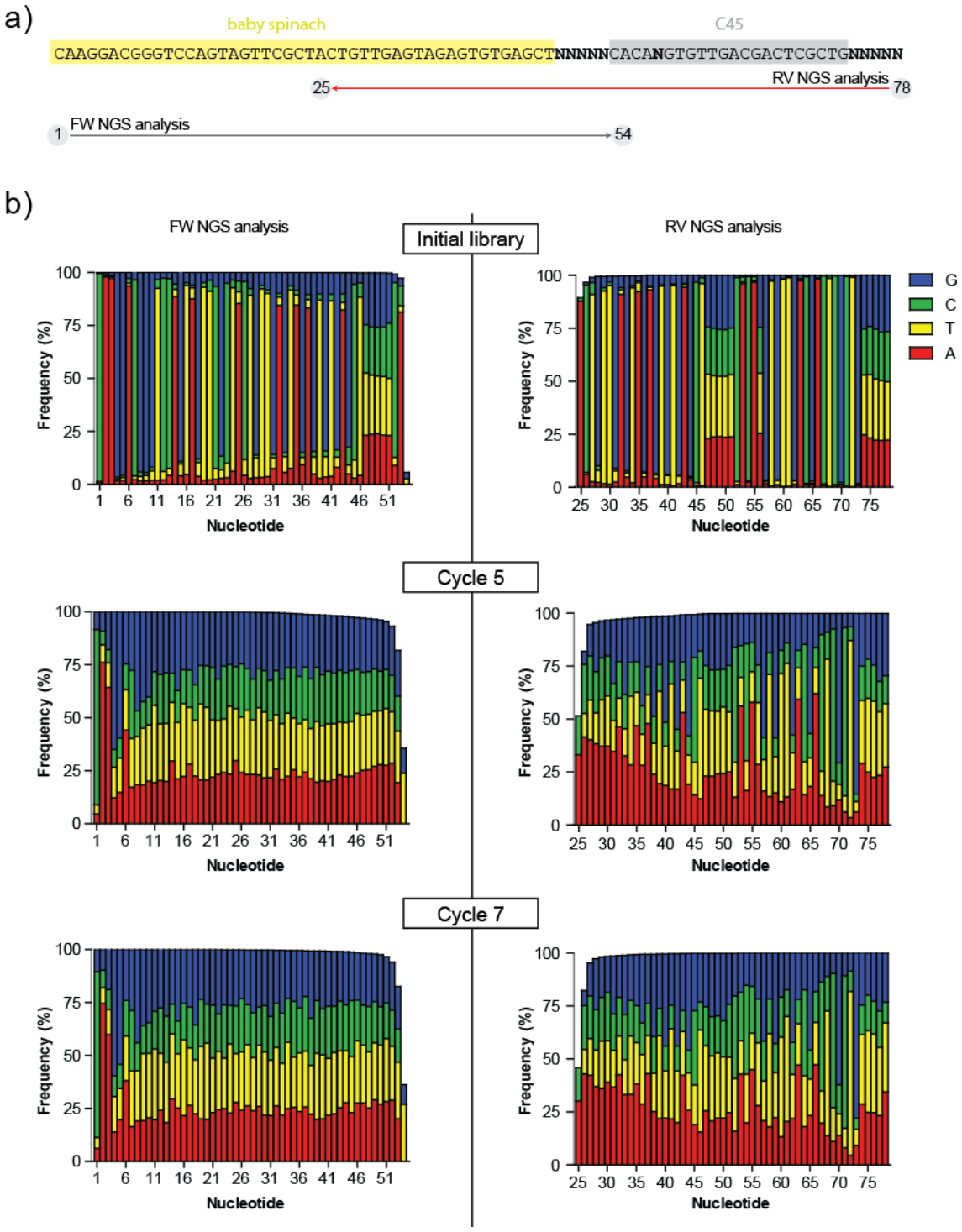
SUPPLEMENTARY FIGURES



Supporting Figure 1 | D1 aptamer interacts linearly with DFHBI and not with eDFHBI
Dose-response fluorescence increase of 10 µM DFHBI (green) and eDFHBI (black) incubated at increasing concentration of D1 aptamer. Due to its stronger affinity to DFHBI, D1 interacts with this dye at <0.02 µM RNA, whereas >0.3 µM RNA (red dotted line) is necessary to increase the fluorescence of eDFHBI.



Supporting Figure 2 | D1 fluorescence quenching assay with mir21 and lysozyme
Fluorescence response of 1 µM D1 at increasing concentration of mir21 (light grey) and lysozyme (dark grey). The fluorescence of D1 is quenched by increasing lysozyme (≥300 nM), whereas mir21 does not affect D1 fluorescence in the 78-5000 nM range. Since miR21 does not seem to quench D1's fluorescence, its SELEX concentration was arbitrarily set to 1000nM. Instead, lysozyme was set to 200 nM to minimize the destabilization of the selected fluorogenic RNAs. Fluorescence assay performed at 1 µM RNA and 10 µM DFHBI in selection buffer (125 mM KCl and 5 mM $MgCl_2$ in 40 mM HEPES pH 7.5).

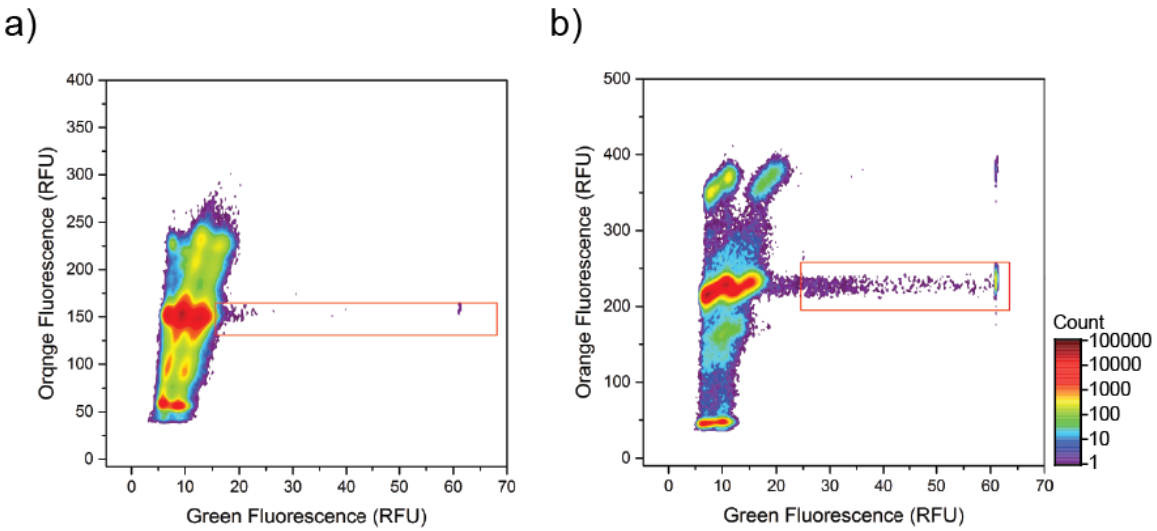


Supporting Figure 3 | Allosteric SELEX leads to randomization in NGS analysis
a, Alignment scheme depicting the library sequence coverage through NGS analysis.

- continues in the next page –

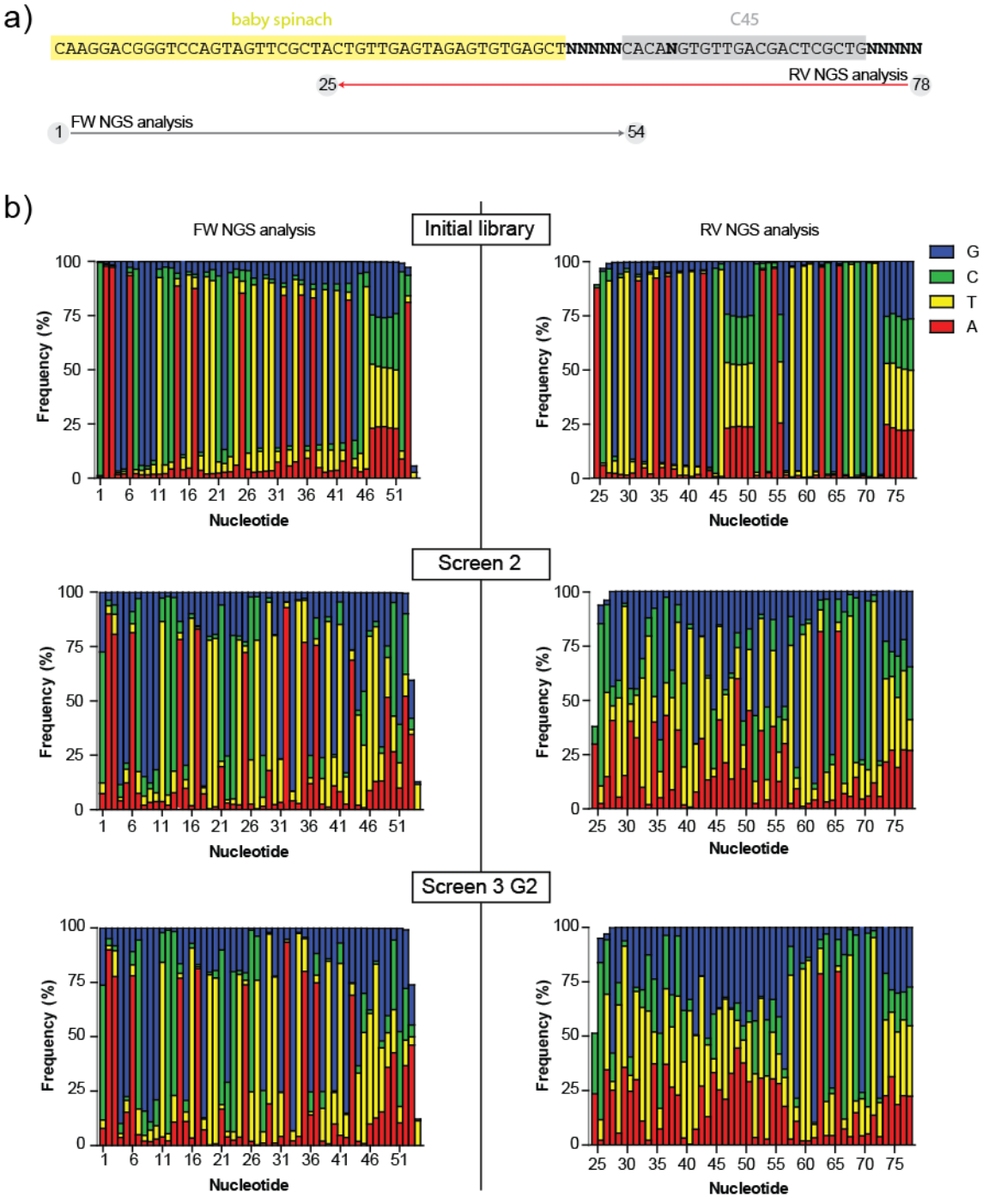
- continuation of the previous page –

Forward NGS analysis covered nucleotides 1-54 of the library, encompassing the Baby Spinach domain (yellow) and part of the random region. Reverse NGS analysis covered nucleotides 78 to 25, encompassing the C45 domain (grey), including all the random nucleotides. The region 25-54, which was sequenced in both directions, was critical for alignment and top-sequence identification. **b,** Nucleotide distribution analysis showing how the library's structural properties are lost throughout the *in vitro* selection due to the appearance of truncated sequences. The structural information of the initial pool is well defined, showing one predominant nucleotide representation in Baby Spinach (1-46 nt) and C45 (52-73 nt), while randomized already at cycle 5 and then cycle 7.



Supporting Figure 4 | µIVC positive screening cycles 1-2 for the selection of FBP responsive RNA light-up devices.

a-b, Fluorescence profile of the droplet emulsion obtained in the first and second positive screening cycle at 2.5 mM FBP. The different populations composing the emulsion were differentiated by DFHBI-1T green fluorescence and Texas-Red orange fluorescence. For the first screening cycle (**a**), the orange fluorescence gate was set to 134-166 RFU, while the green fluorescence gate was 16-68 RFU. In the second screening cycle (**b**), the orange fluorescence gate was set to 195-251 RFU, while the green fluorescence gate was 26-68 RFU.

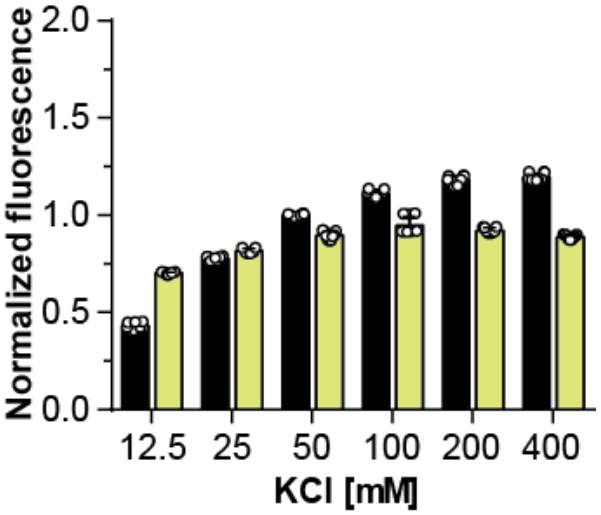


Supporting Figure 5 | µIVC maintains the structural identity of the chimeric library. **a,** Alignment scheme depicting the library sequence coverage through NGS analysis.

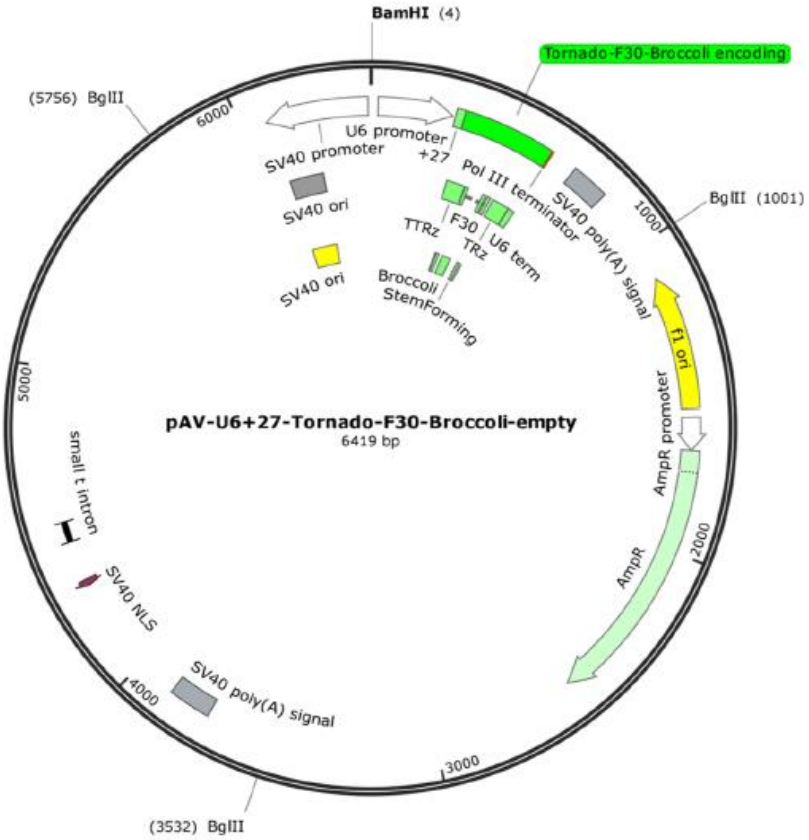
- continues in the next page –

– continuation of previous page –

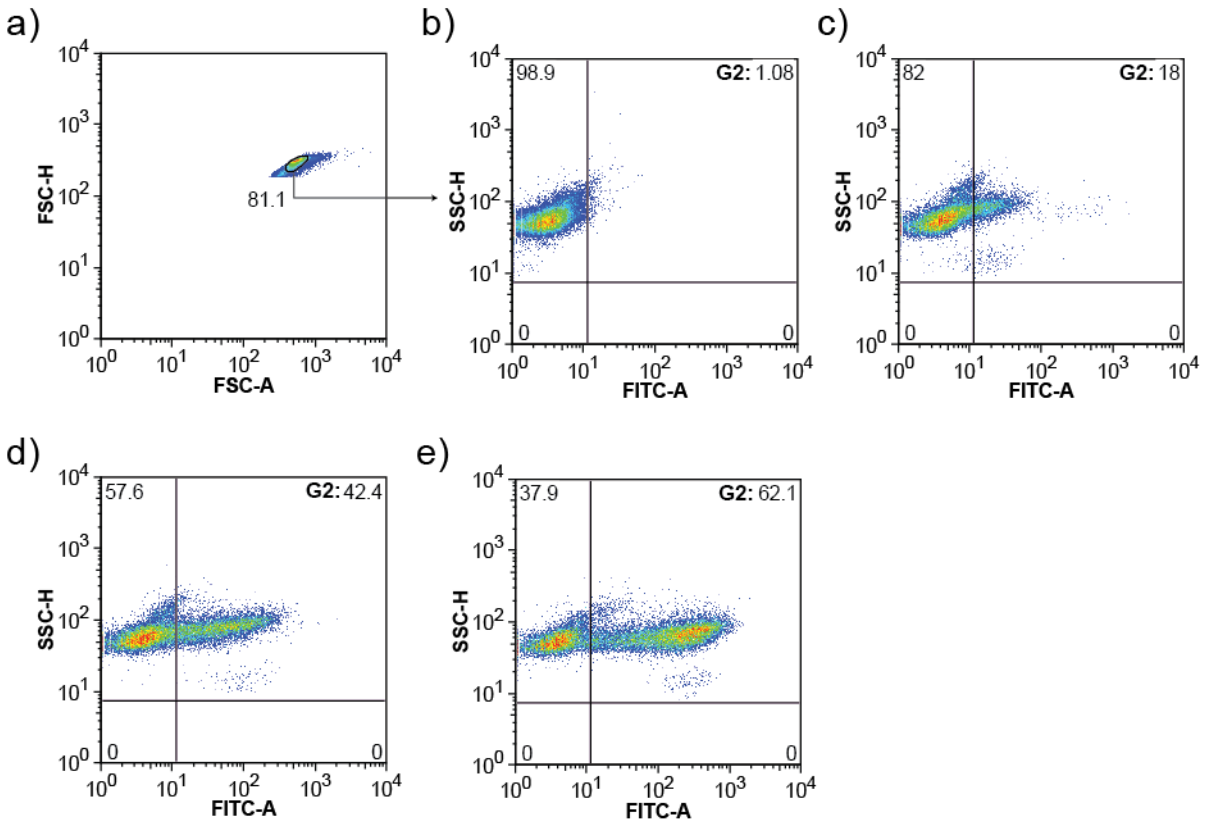
Forward NGS analysis covered nucleotides 1-54 of the library, encompassing spinach baby domain (yellow) and part of the random region. Reverse NGS analysis covered nucleotides 78 to 25, encompassing the C45 domain (grey), including all the random nucleotides. The region 25-54, which was sequenced in both directions, was critical for alignment and top-sequence identification. **b,** Nucleotide distribution analysis showing how the library's structural properties are preserved. The structural information of the initial pool is preserved throughout the screening, predominantly in the region of Baby Spinach (1-46 nt). C45 region is preserved to a lesser extent, showing in particular high variability in the transducer region (47-56 nt) due to the enrichment, as shown in **Tables 2-3**, of a truncated family where this region is lacking.



Supporting Figure 6| K^+ ions stabilize D1 G-quadruplex
Fluorescence readout of D1 aptamer control measured in the absence (black) and presence of 40 mM FBP (yellow) at increasing concentration of KCl. Values were normalized by setting to 1 the background fluorescence at 50 mM KCl, The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9 and 30 mM $MgCl_2$.

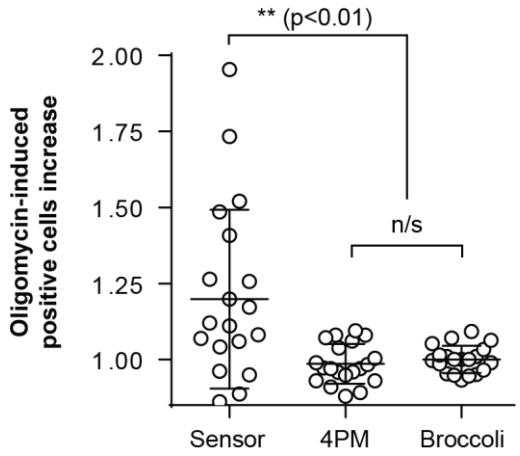


Supporting Figure 7| Plasmid map of pAV-U6+27-Tornado-F30-Broccoli-empty.
The sequence information and plasmid specs were published by Litke *et al.*[135] The plasmid map was generated with SnapGene Viewer.

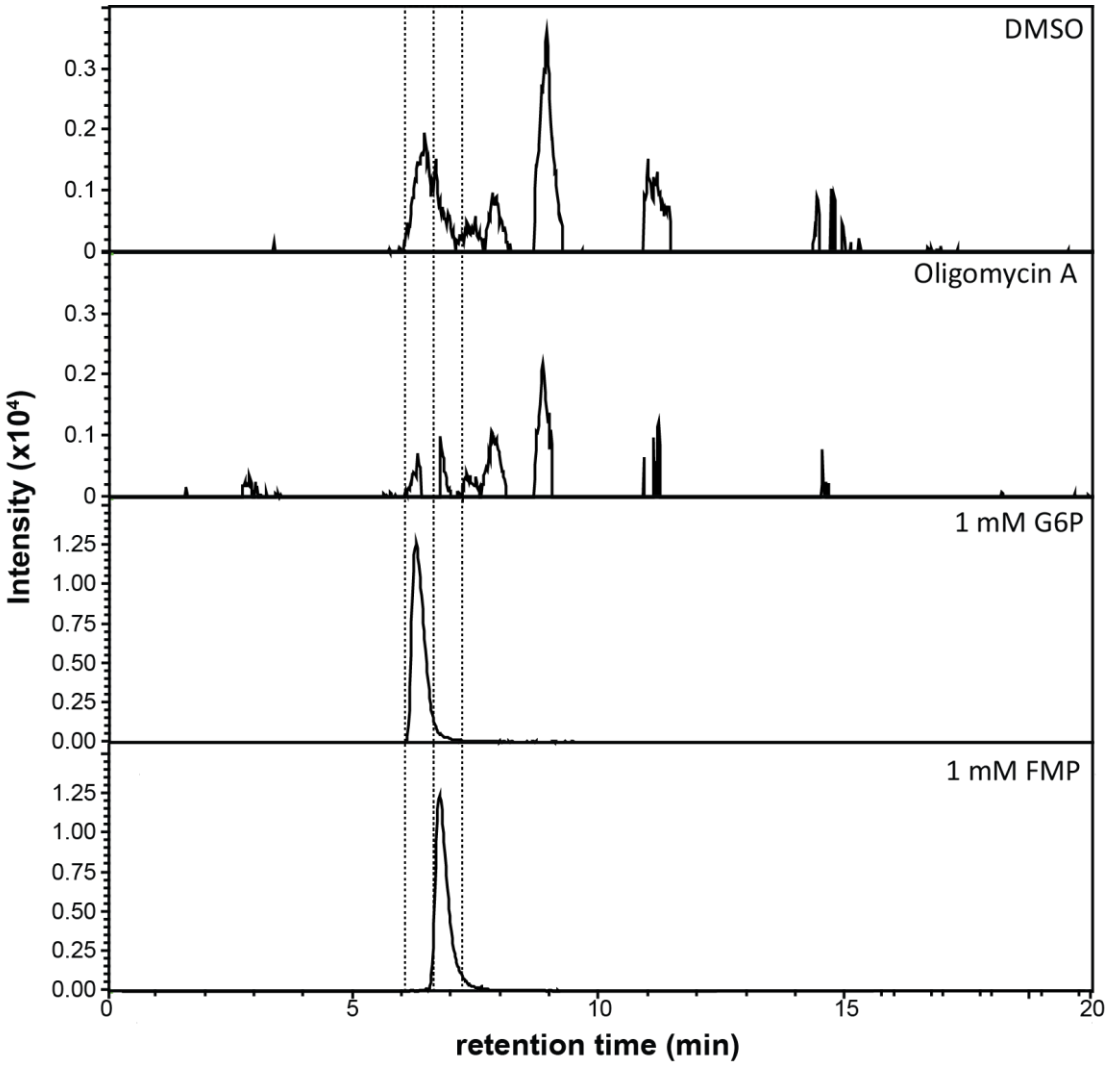


Supporting Figure 8| Gating strategy to identify positive cells

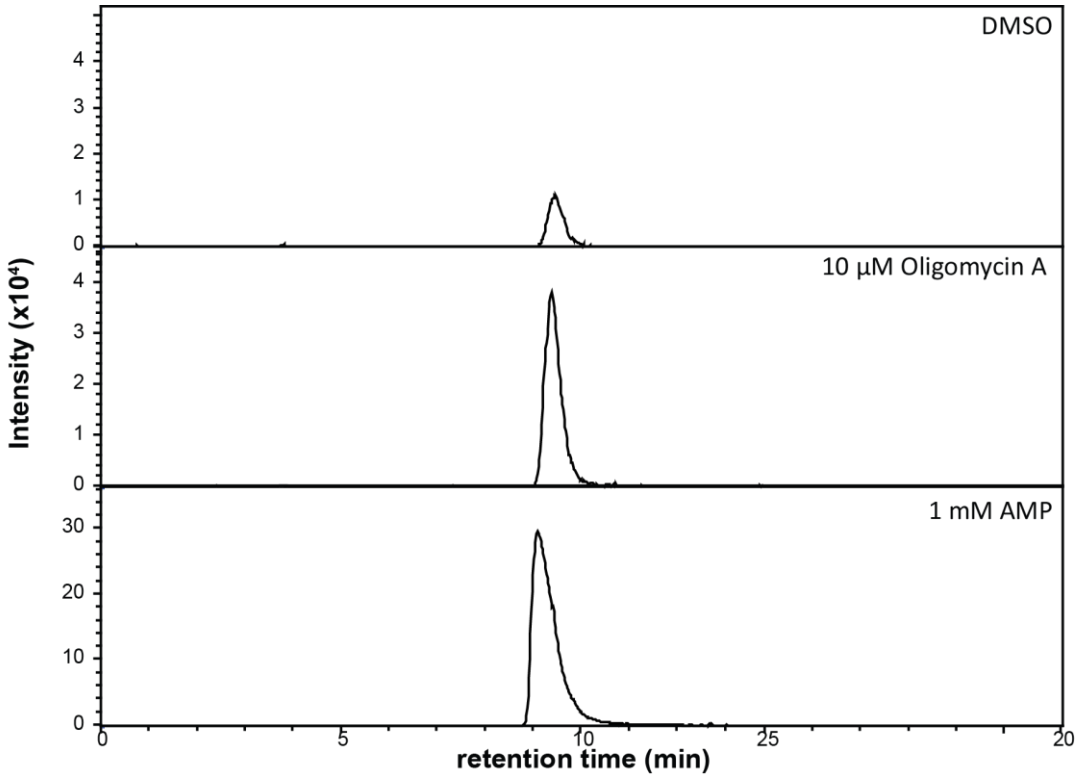
a, Singlet population was gated for untransfected cell sample using forward scatter height (FSC-H) vs. forward scatter area (FSC-A) density plot. **b,** The singlet untransfected population was then plotted using side scatter height (FSC-H) vs. Fluorescein isothiocyanate area (FITC-A) to identify the 1% positive cell gate (G2). The 1% gate was applied to s81 (**c**), 4PM (**d**), and Broccoli (**e**) to obtain the positive cell percentile (G2), which was plotted in single parameter histograms (**Figure 28**).



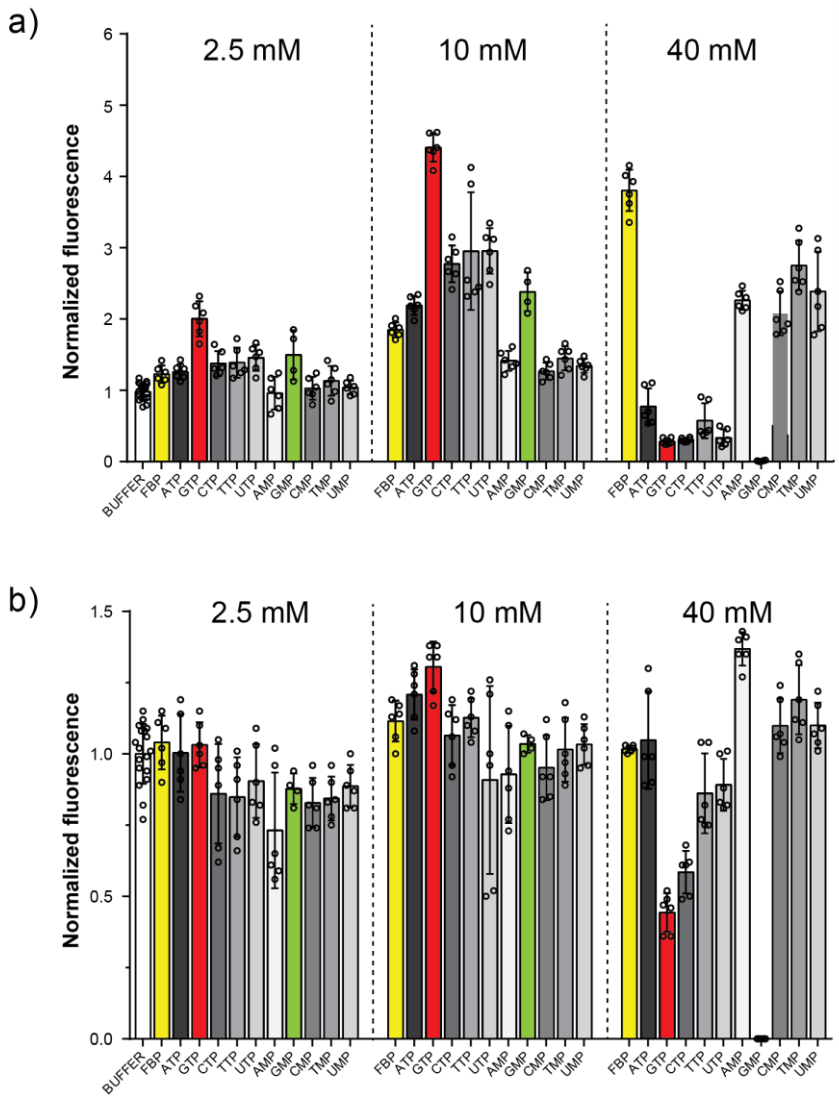
Supporting Figure 9| Double-blind confirmation of the correlation between 81 and oligomycin
The experimental setup presented in **Figure 30** was repeated in a double-blind setting (n=2) and included with the previous n=3 repetition, increasing the p-value of the test from 0.0175 (*, **Figure 30 b**) to 0.006 (**). The p-values were calculated using the Mann-Whitney non-parametric test.



Supporting Figure 10| FMP and G6P decrease in oligomycin-treated HEK293T cells
LC-MS analysis of intracellular glucose 6-phosphate (G6P) and fructose 6-phosphate (FMP) levels in HEK293T cells treated with oligomycin A or DMSO solvent vehicle (DMSO). A parallel LC-MS run was performed with 1 mM of each metabolite to identify their retention time and pick area, which is highlighted (dotted lines).



Supporting Figure 11| AMP robustly increases in oligomycin-treated HEK293T cells
LC-MS analysis of intracellular adenosine monophosphate (AMP) levels in HEK293T cells treated with oligomycin A or DMSO solvent vehicle (DMSO). A parallel LC-MS run was performed with 1 mM of each metabolite to identify the metabolite's retention time and pick area.



Supporting Figure 12| Fluorescence response of s81 and D1 aptamer to NTPs and nMPs
The fluorescence response of s81 (**a**) and D1 aptamer control (**b**) to a panel of triphosphate and monophosphate nucleotides, at the given concentrations, was normalized to the background fluorescence of each RNA (buffer) and was set to 1. The final specificity assay shown in **Figure 33** was obtained by dividing s81's response (**a**) to the corresponding value of D1 (**b**) for each condition. The RNAs response to FBP (yellow), GTP (red), and GMP (green) was color-coded to be highlighted. The fluorescence readout was measured at 1 µM RNA and 10 µM DFHBI in 40 mM Tris-HCl pH 7.9, 50 mM KCl, 30 mM $MgCl_2$.

8 REFERENCES

1. Nachmanson, D., Ochoa, S. & Lipmann, F.A. Otto Meyerhof (1884-1951): A Biographical Memoir. *National Academy of Science* (1960).
2. Schirmer, H. & Gromer, S. Meyerhof in Heidelberg - The start of cell biology. *Web Archive* (2001).
3. Warburg, O. On the origin of cancer cells. *Science* **123**, 309-14 (1956).
4. Fothergill-Gilmore, L.A. & Michels, P.A. Evolution of glycolysis. *Prog Biophys Mol Biol* **59**, 105-235 (1993).
5. Li, X.B., Gu, J.D. & Zhou, Q.H. Review of aerobic glycolysis and its key enzymes - new targets for lung cancer therapy. *Thorac Cancer* **6**, 17-24 (2015).
6. Berg, J.M., Tymoczko, J.L. & Stryer, L. Biochemistry. 5th edition. *W. H. Freeman and Company* (2002).
7. Garrett, R.H. & Grisham, C.M. Biochemistry. 4th edition. *Brooks/Cole, Cengage Learning* (2010).
8. Walsh, C.T., Tu, B.P. & Tang, Y. Eight Kinetically Stable but Thermodynamically Activated Molecules that Power Cell Metabolism. *Chem Rev* **118**, 1460-1494 (2018).
9. TeSlaa, T. & Teitell, M.A. Techniques to monitor glycolysis. *Methods Enzymol* **542**, 91-114 (2014).
10. Lee, M. & Yoon, J.H. Metabolic interplay between glycolysis and mitochondrial oxidation: The reverse Warburg effect and its therapeutic implication. *World J Biol Chem* **6**, 148-61 (2015).
11. Locasale, J.W. Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nat Rev Cancer* **13**, 572-83 (2013).
12. De Bock, K. et al. Role of PFKFB3-driven glycolysis in vessel sprouting. *Cell* **154**, 651-63 (2013).

13. Buck, M.D., Sowell, R.T., Kaech, S.M. & Pearce, E.L. Metabolic Instruction of Immunity. *Cell* **169**, 570-586 (2017).

14. Yu, P. et al. FGF-dependent metabolic control of vascular development. *Nature* **545**, 224-228 (2017).

15. Tanner, L.B. et al. Four Key Steps Control Glycolytic Flux in Mammalian Cells. *Cell Syst* **7**, 49-62 e8 (2018).

16. Fan, J. et al. Glutamine-driven oxidative phosphorylation is a major ATP source in transformed mammalian cells in both normoxia and hypoxia. *Mol Syst Biol* **9**, 712 (2013).

17. DeBerardinis, R.J. & Chandel, N.S. Fundamentals of cancer metabolism. *Sci Adv* **2**, e1600200 (2016).

18. Vander Heiden, M.G., Cantley, L.C. & Thompson, C.B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029-33 (2009).

19. Liberti, M.V. & Locasale, J.W. The Warburg Effect: How Does it Benefit Cancer Cells? (vol 41, pg 211, 2016). *Trends in Biochemical Sciences* **41**, 287-287 (2016).

20. Guppy, M., Leedman, P., Zu, X. & Russell, V. Contribution by different fuels and metabolic pathways to the total ATP turnover of proliferating MCF-7 breast cancer cells. *Biochem J* **364**, 309-15 (2002).

21. Martin, M. et al. Energetic and morphological plasticity of C6 glioma cells grown on 3-D support; effect of transient glutamine deprivation. *J Bioenerg Biomembr* **30**, 565-78 (1998).

22. Smolkova, K. et al. Mitochondrial bioenergetic adaptations of breast cancer cells to aglycemia and hypoxia. *J Bioenerg Biomembr* **42**, 55-67 (2010).

23. Sonveaux, P. et al. Targeting lactate-fueled respiration selectively kills hypoxic tumor cells in mice. *J Clin Invest* **118**, 3930-42 (2008).

24. Takhaveev, V. & Heinemann, M. Metabolic heterogeneity in clonal microbial populations. *Curr Opin Microbiol* **45**, 30-38 (2018).

25. Song, M. & Giovannucci, E.L. Cancer risk: many factors contribute. *Science* **347**, 728-9 (2015).

26. Moreno-Sanchez, R., Rodriguez-Enriquez, S., Marin-Hernandez, A. & Saavedra, E. Energy metabolism in tumor cells. *FEBS J* **274**, 1393-418 (2007).

27. Shin, Y.S. et al. Quantitative assessments of glycolysis from single cells. *Technology (Singap World Sci)* **3**, 172-178 (2015).

28. van Heerden, J.H. et al. Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. *Science* **343**, 1245114 (2014).

29. Elowitz, M.B., Levine, A.J., Siggia, E.D. & Swain, P.S. Stochastic gene expression in a single cell. *Science* **297**, 1183-6 (2002).

30. Campbell, K., Vowinckel, J. & Ralser, M. Cell-to-cell heterogeneity emerges as consequence of metabolic cooperation in a synthetic yeast community. *Biotechnol J* **11**, 1169-78 (2016).

31. Papagiannakis, A., Niebel, B., Wit, E.C. & Heinemann, M. Autonomous Metabolic Oscillations Robustly Gate the Early and Late Cell Cycle. *Mol Cell* **65**, 285-295 (2017).

32. Kamei, Y., Tamada, Y., Nakayama, Y., Fukusaki, E. & Mukai, Y. Changes in transcription and metabolism during the early stage of replicative cellular senescence in budding yeast. *J Biol Chem* **289**, 32081-93 (2014).

33. Winer, L.S.P. & Wu, M. Rapid Analysis of Glycolytic and Oxidative Substrate Flux of Cancer Cells in a Microplate. *Plos One* **9**(2014).

34. Zenobi, R. Single-cell metabolomics: analytical and biological perspectives. *Science* **342**, 1243259 (2013).

35. Duncan, K.D., Fyrestam, J. & Lanekoff, I. Advances in mass spectrometry based single-cell metabolomics. *Analyst* **144**, 782-793 (2019).

36. Bermejo, C., Ewald, J.C., Lanquar, V., Jones, A.M. & Frommer, W.B. In vivo biochemistry: quantifying ion and metabolite levels in individual cells or cultures of yeast. *Biochem J* **438**, 1-10 (2011).

37. Chudakov, D.M., Matz, M.V., Lukyanov, S. & Lukyanov, K.A. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* **90**, 1103-63 (2010).

38. Su, Y. & Hammond, M.C. RNA-based fluorescent biosensors for live cell imaging of small molecules and RNAs. *Curr Opin Biotechnol* **63**, 157-166 (2020).

39. Siedler, S. et al. SoxR as a single-cell biosensor for NADPH-consuming enzymes in Escherichia coli. *ACS Synth Biol* **3**, 41-7 (2014).

40. Binder, S. et al. A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol* **13**, R40 (2012).

41. Raman, S., Rogers, J.K., Taylor, N.D. & Church, G.M. Evolution-guided optimization of biosynthetic pathways. *Proc Natl Acad Sci U S A* **111**, 17803-8 (2014).

42. Michener, J.K., Thodey, K., Liang, J.C. & Smolke, C.D. Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab Eng* **14**, 212-22 (2012).

43. Mor, I., Cheung, E.C. & Vousden, K.H. Control of glycolysis through regulation of PFK1: old friends and recent additions. *Cold Spring Harb Symp Quant Biol* **76**, 211-6 (2011).

44. Huberts, D.H., Niebel, B. & Heinemann, M. A flux-sensing mechanism could regulate the switch between respiration and fermentation. *FEMS Yeast Res* **12**, 118-28 (2012).

45. Bosch, D. et al. Characterization of glucose transport mutants of Saccharomyces cerevisiae during a nutritional upshift reveals a correlation between metabolite levels and glycolytic flux. *FEMS Yeast Res* **8**, 10-25 (2008).

46. Monteiro, F. et al. Measuring glycolytic flux in single yeast cells with an orthogonal synthetic biosensor. *Mol Syst Biol* **15**, e9071 (2019).

47. Kochanowski, K. et al. Functioning of a metabolic flux sensor in Escherichia coli. *Proc Natl Acad Sci U S A* **110**, 1130-5 (2013).

48. Kotte, O., Volkmer, B., Radzikowski, J.L. & Heinemann, M. Phenotypic bistability in Escherichia coli's central carbon metabolism. *Mol Syst Biol* **10**, 736 (2014).

49. Shimomura, O., Johnson, F.H. & Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from Luminous Hydromedusan, Aequorea. *Journal of Cellular and Comparative Physiology* **59**, 223-& (1962).

50. Barrick, J.E. & Breaker, R.R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* **8**, R239 (2007).

51. Carpenter, A.C., Paulsen, I.T. & Williams, T.C. Blueprints for Biosensors: Design, Limitations, and Applications. *Genes (Basel)* **9**(2018).

52. Hou, B.H. et al. Optical sensors for monitoring dynamic changes of intracellular metabolite levels in mammalian cells. *Nat Protoc* **6**, 1818-33 (2011).

53. San Martin, A. et al. Imaging mitochondrial flux in single cells with a FRET sensor for pyruvate. *PLoS One* **9**, e85780 (2014).

54. Specht, E.A., Braselmann, E. & Palmer, A.E. A Critical and Comparative Review of Fluorescent Tools for Live-Cell Imaging. *Annual Review of Physiology, Vol 79* **79**, 93-117 (2017).

55. Wu, G. et al. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol* **34**, 652-664 (2016).

56. Borkowski, O. et al. Cell-free prediction of protein expression costs for growing cells. *Nat Commun* **9**, 1457 (2018).

57. Ceroni, F., Algar, R., Stan, G.B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat Methods* **12**, 415-8 (2015).

58. Famulok, M. & Mayer, G. Aptamers and SELEX in Chemistry & Biology. *Chemistry & Biology* **21**, 1055-1058 (2014).

59. Ali, M.H., Elsherbiny, M.E. & Emara, M. Updates on Aptamer Research. *International Journal of Molecular Sciences* **20**(2019).

60. Pfeiffer, F. Click-SELEX enables the selection of Δ9-tetrahydrocannabinol-binding nucleic acids. *PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn* (2018).

61. Weber, A.M. et al. A blue light receptor that mediates RNA binding and translational regulation. *Nat Chem Biol* **15**, 1085-1092 (2019).

62. Lennarz, S., Heider, E., Blind, M. & Mayer, G. An aptamer to the MAP kinase insert region. *ACS Chem Biol* **10**, 320-7 (2015).

63. Win, M.N., Klein, J.S. & Smolke, C.D. Codeine-binding RNA aptamers and rapid determination of their binding constants using a direct coupling surface plasmon resonance assay. *Nucleic Acids Research* **34**, 5670-5682 (2006).

64. Koizumi, M. & Breaker, R.R. Molecular recognition of cAMP by an RNA aptamer. *Biochemistry* **39**, 8983-92 (2000).

65. Ciesiolka, J., Gorski, J. & Yarus, M. Selection of an Rna Domain That Binds Zn2+. *Rna* **1**, 538-550 (1995).

66. Hofmann, H.P., Limmer, S., Hornung, V. & Sprinzl, M. Ni2+-binding RNA motifs with an asymmetric purine-rich internal loop and a G-A base pair. *RNA (New York, N.Y.)* **3**, 1289-300 (1997).

67. Mi, J. et al. In vivo selection of tumor-targeting RNA motifs. *Nat Chem Biol* **6**, 22-4 (2010).

68. Schroeder, R. Soups & SELEX for the origin of life. *RNA* **21**, 729-32 (2015).

69. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-10 (1990).

70. Gold, L. SELEX: How It Happened and Where It will Go. *Journal of Molecular Evolution* **81**, 140-143 (2015).

71. Pfeiffer, F. & Mayer, G. Selection and Biosensor Application of Aptamers for Small Molecules. *Front Chem* **4**, 25 (2016).

72. Cho, E.J., Lee, J.-W. & Ellington, A.D. Applications of aptamers as sensors. *Annual review of analytical chemistry (Palo Alto, Calif.)* **2**, 241-64 (2009).

73. Grate, D. & Wilson, C. Laser-mediated, site-specific inactivation of RNA transcripts. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6131-6136 (1999).

74. Holeman, L.A., Robinson, S.L., Szostak, J.W. & Wilson, C. Isolation and characterization of fluorophore-binding RNA aptamers. *Folding & Design* **3**, 423-431 (1998).

75. Babendure, J.R., Adams, S.R. & Tsien, R.Y. Aptamers switch on fluorescence of triphenylmethane dyes. *J Am Chem Soc* **125**, 14716-7 (2003).

76. Arora, A., Sunbul, M. & Jaschke, A. Dual-colour imaging of RNAs using quencher- and fluorophore-binding aptamers. *Nucleic Acids Res* **43**, e144 (2015).

77. Paige, J.S., Wu, K.Y. & Jaffrey, S.R. RNA mimics of green fluorescent protein. *Science* **333**, 642-6 (2011).

78. Pothoulakis, G., Ceroni, F., Reeve, B. & Ellis, T. The spinach RNA aptamer as a characterization tool for synthetic biology. *ACS Synth Biol* **3**, 182-7 (2014).

79. Zhang, J. et al. Tandem Spinach Array for mRNA Imaging in Living Bacterial Cells. *Scientific Reports* **5**, 17295 (2015).

80. Guet, D. et al. Combining Spinach-tagged RNA and gene localization to image gene expression in live yeast. *Nat Commun* **6**(2015).

81. Tsvetkova, I.B., Yi, G., Yi, Y., Kao, C.C. & Dragnea, B.G. Segmented GFP-like aptamer probes for functional imaging of viral genome trafficking. *Virus Research* **210**, 291-297 (2015).

82. Burch, B.D., Garrido, C. & Margolis, D.M. Detection of human immunodeficiency virus RNAs in living cells using Spinach RNA aptamers. *Virus Res* **228**, 141-146 (2017).

83. Strack, R.L., Disney, M.D. & Jaffrey, S.R. A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat Methods* **10**, 1219-24 (2013).

84. Guzman-Zapata, D. et al. mRNA imaging in the chloroplast of Chlamydomonas reinhardtii using the light-up aptamer Spinach. *J Biotechnol* (2017).

85. Hofer, K., Langejurgen, L.V. & Jaschke, A. Universal aptamer-based real-time monitoring of enzymatic RNA synthesis. *J Am Chem Soc* **135**, 13692-4 (2013).

86. Chizzolini, F., Forlin, M., Cecchi, D. & Mansy, S.S. Gene position more strongly influences cell-free protein expression from operons than T7 transcriptional promoter strength. *ACS Synth Biol* **3**, 363-71 (2014).

87. Chizzolini, F. et al. Cell-Free Translation Is More Variable than Transcription. *ACS Synth Biol* **6**, 638-647 (2017).

88. Li, B., Ellington, A.D. & Chen, X. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Res* **39**, e110 (2011).

89. Bhadra, S. & Ellington, A.D. Design and application of cotranscriptional non-enzymatic RNA circuits and signal transducers. *Nucleic Acids Res* **42**, e58 (2014).

90. Svensen, N. & Jaffrey, S.R. Fluorescent RNA Aptamers as a Tool to Study RNA-Modifying Enzymes. *Cell Chem Biol* **23**, 415-25 (2016).

91. Abatemarco, J. et al. RNA-aptamers-in-droplets (RAPID) high-throughput screening for secretory phenotypes. *Nat Commun* **8**, 332 (2017).

92. Han, K.Y., Leslie, B.J., Fei, J., Zhang, J. & Ha, T. Understanding the photophysics of the spinach-DFHBI RNA aptamer-fluorogen complex to improve live-cell RNA imaging. *J Am Chem Soc* **135**, 19033-8 (2013).

93. Lee, H.-W., Robinson, S.G., Bandyopadhyay, S., Mitchell, R.H. & Sen, D. Reversible photo-regulation of a hammerhead ribozyme using a diffusible effector. *Journal of molecular biology* **371**, 1163-73 (2007).

94. Warner, K.D. et al. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat Struct Mol Biol* **21**, 658-63 (2014).

95. Filonov, G.S., Moon, J.D., Svensen, N. & Jaffrey, S.R. Broccoli: Rapid Selection of an RNA Mimic of Green Fluorescent Protein by Fluorescence-Based Selection and Directed Evolution. *Journal of the American Chemical Society* **136**, 16299-16308 (2014).

96. Warner, K.D. et al. A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nat Chem Biol* (2017).

97. Song, W. et al. Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex. *Nat Chem Biol* (2017).

98. Bouhedda, F. et al. A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells. *Nat Chem Biol* **16**, 69-76 (2020).

99. Steinmetzger, C., Palanisamy, N., Gore, K.R. & Hobartner, C. A Multicolor Large Stokes Shift Fluorogen-Activating RNA Aptamer with Cationic Chromophores. *Chemistry* **25**, 1931-1935 (2019).

100. Dolgosheina, E.V. et al. RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS Chem Biol* **9**, 2412-20 (2014).

101. Trachman, R.J., 3rd et al. Crystal Structures of the Mango-II RNA Aptamer Reveal Heterogeneous Fluorophore Binding and Guide Engineering of Variants with Improved Selectivity and Brightness. *Biochemistry* **57**, 3544-3548 (2018).

102. Trachman, R.J., 3rd et al. Structure and functional reselection of the Mango-III fluorogenic RNA aptamer. *Nat Chem Biol* **15**, 472-479 (2019).

103. Wirth, R., Gao, P., Nienhaus, G.U., Sunbul, M. & Jaschke, A. SiRA: A Silicon Rhodamine-Binding Aptamer for Live-Cell Super-Resolution RNA Imaging. *J Am Chem Soc* (2019).

104. Huang, H. et al. A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nat Chem Biol* **10**, 686-91 (2014).

105. Ouellet, J. RNA Fluorescence with Light-Up Aptamers. *Front Chem* **4**, 29 (2016).

106. Remington, S.J. Green fluorescent protein: a perspective. *Protein Sci* **20**, 1509-19 (2011).

107. Paul, B.K. & Guchhait, N. Looking at the Green Fluorescent Protein (GFP) chromophore from a different perspective: a computational insight. *Spectrochim Acta A Mol Biomol Spectrosc* **103**, 295-303 (2013).

108. Wang, P. et al. Photochemical properties of Spinach and its use in selective imaging. *Chemical Science* **4**, 2865-2873 (2013).

109. Song, W., Strack, R.L., Svensen, N. & Jaffrey, S.R. Plug-and-play fluorophores extend the spectral properties of Spinach. *J Am Chem Soc* **136**, 1198-201 (2014).

110. Strack, R.L. & Jaffrey, S.R. Live-cell imaging of mammalian RNAs with Spinach2. *Methods Enzymol* **550**, 129-46 (2015).

111. Filonov, G.S., Moon, J.D., Svensen, N. & Jaffrey, S.R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J Am Chem Soc* **136**, 16299-308 (2014).

112. Filonov, G.S., Kam, C.W., Song, W. & Jaffrey, S.R. In-gel imaging of RNA processing using broccoli reveals optimal aptamer expression strategies. *Chem Biol* **22**, 649-60 (2015).

113. Filonov, G.S. & Jaffrey, S.R. RNA Imaging with Dimeric Broccoli in Live Bacterial and Mammalian Cells. *Curr Protoc Chem Biol* **8**, 1-28 (2016).

114. Autour, A., Westhof, E. & Ryckelynck, M. iSpinach: a fluorogenic RNA aptamer optimized for in vitro applications. *Nucleic Acids Res* (2016).

115. Okuda, M., Fourmy, D. & Yoshizawa, S. Use of Baby Spinach and Broccoli for imaging of structured cellular RNAs. *Nucleic Acids Res* (2016).

116. Baret, J.C. et al. Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip* **9**, 1850-8 (2009).

117. Furuhata, Y. et al. Programmable RNA detection with a fluorescent RNA aptamer using optimized three-way junction formation. *RNA* **25**, 590-599 (2019).

118. Soni, R., Sharma, D., Krishna, A.M., Sathiri, J. & Sharma, A. A highly efficient Baby Spinach-based minimal modified sensor (BSMS) for nucleic acid analysis. *Org Biomol Chem* **17**, 7222-7227 (2019).

119. Ying, Z.M., Wu, Z., Tu, B., Tan, W. & Jiang, J.H. Genetically Encoded Fluorescent RNA Sensor for Ratiometric Imaging of MicroRNA in Living Tumor Cells. *J Am Chem Soc* **139**, 9779-9782 (2017).

120. Bhadra, S. & Ellington, A.D. A Spinach molecular beacon triggered by strand displacement. *RNA* **20**, 1183-94 (2014).

121. Strack, R.L. & Jaffrey, S.R. New approaches for sensing metabolites and proteins in live cells using RNA. *Curr Opin Chem Biol* **17**, 651-5 (2013).

122. Paige, J.S., Nguyen-Duc, T., Song, W. & Jaffrey, S.R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194 (2012).

123. Kellenberger, C.A., Wilson, S.C., Sales-Lee, J. & Hammond, M.C. RNA-based fluorescent biosensors for live cell imaging of second messengers cyclic di-GMP and cyclic AMP-GMP. *J Am Chem Soc* **135**, 4906-9 (2013).

124. Kellenberger, C.A., Chen, C., Whiteley, A.T., Portnoy, D.A. & Hammond, M.C. RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messenger Cyclic di-AMP. *J Am Chem Soc* **137**, 6432-5 (2015).

125. You, M., Litke, J.L. & Jaffrey, S.R. Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proc Natl Acad Sci U S A* **112**, E2756-65 (2015).

126. Bose, D., Su, Y., Marcus, A., Raulet, D.H. & Hammond, M.C. An RNA-Based Fluorescent Biosensor for High-Throughput Analysis of the cGAS-cGAMP-STING Pathway. *Cell Chem Biol* (2016).

127. Ketterer, S., Gladis, L., Kozica, A. & Meier, M. Engineering and characterization of fluorogenic glycine riboswitches. *Nucleic Acids Res* **44**, 5983-92 (2016).

128. Wang, X.C., Wilson, S.C. & Hammond, M.C. Next-generation RNA-based fluorescent biosensors enable anaerobic detection of cyclic di-GMP. *Nucleic Acids Res* **44**, e139 (2016).

129. Porter, E.B., Polaski, J.T., Morck, M.M. & Batey, R.T. Recurrent RNA motifs as scaffolds for genetically encodable small-molecule biosensors. *Nat Chem Biol* (2017).

130. Autour, A., Bouhedda, F., Cubi, R. & Ryckelynck, M. Optimization of fluorogenic RNA-based biosensors using droplet-based microfluidic ultrahigh-throughput screening. *Methods* (2019).

131. You, M., Litke, J.L., Wu, R. & Jaffrey, S.R. Detection of Low-Abundance Metabolites in Live Cells Using an RNA Integrator. *Cell Chem Biol* (2019).

132. Kim, H. & Jaffrey, S.R. A Fluorogenic RNA-Based Sensor Activated by Metabolite-Induced RNA Dimerization. *Cell Chem Biol* **26**, 1725-1731 e6 (2019).

133. Porter, E.B., Polaski, J.T., Morck, M.M. & Batey, R.T. Recurrent RNA motifs as scaffolds for genetically encodable small-molecule biosensors. *Nat Chem Biol* **13**, 295-301 (2017).

134. Su, Y., Hickey, S.F., Keyser, S.G. & Hammond, M.C. In Vitro and In Vivo Enzyme Activity Screening via RNA-Based Fluorescent Biosensors for S-Adenosyl-l-homocysteine (SAH). *J Am Chem Soc* **138**, 7040-7 (2016).

135. Litke, J.L. & Jaffrey, S.R. Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat Biotechnol* (2019).

136. Tizei, P.A., Csibra, E., Torres, L. & Pinheiro, V.B. Selection platforms for directed evolution in synthetic biology. *Biochem Soc Trans* **44**, 1165-75 (2016).

137. Hallberg, Z.F., Su, Y., Kitto, R.Z. & Hammond, M.C. Engineering and In Vivo Applications of Riboswitches. *Annu Rev Biochem* **86**, 515-539 (2017).

138. Nutiu, R. & Li, Y. In Vitro Selection of Structure-Switching Signaling Aptamers. *Angewandte Chemie International Edition* **44**, 1061-1065 (2005).

139. Stoltenburg, R., Nikolaus, N. & Strehlitz, B. Capture-SELEX: Selection of DNA Aptamers for Aminoglycoside Antibiotics. *J Anal Methods Chem* **2012**, 415697 (2012).

140. Boussebayle, A., Groher, F. & Suess, B. Corrigendum to "RNA-based Capture-SELEX for the selection of small molecule-binding aptamers" [Methods 161 (2019) 10-15]. *Methods* (2019).

141. Ryckelynck, M. et al. Using droplet-based microfluidics to improve the catalytic properties of RNA under multiple-turnover conditions. *RNA* **21**, 458-69 (2015).

142. Davis, J.H. & Szostak, J.W. Isolation of high-affinity GTP aptamers from partially structured RNA libraries. *Proc Natl Acad Sci U S A* **99**, 11616-21 (2002).

143. Breuers, S., Bryant, L.L., Legen, T. & Mayer, G. Robotic assisted generation of 2'-deoxy-2'-fluoro-modifed RNA aptamers - High performance enabling strategies in aptamer selection. *Methods* **161**, 3-9 (2019).

144. Pfeiffer, F. et al. Identification and characterization of nucleobase-modified aptamers by click-SELEX. *Nature Protocols* **13**, 1153-1180 (2018).

145. Geiger, A., Burgstaller, P., vonderEltz, H., Roeder, A. & Famulok, M. RNA aptamers that bind L-arginine with sub-micromolar dissociation constants and high enantioselectivity. *Nucleic Acids Research* **24**, 1029-1036 (1996).

146. Thiel, W.H. et al. Rapid identification of cell-specific, internalizing RNA aptamers with bioinformatics analyses of a cell-based aptamer selection. *PLoS One* **7**, e43836 (2012).

147. Ageely, E.A., Kartje, Z.J., Rohilla, K.J., Barkau, C.L. & Gagnon, K.T. Quadruplex-Flanking Stem Structures Modulate the Stability and Metal Ion Preferences of RNA Mimics of GFP. *ACS Chem Biol* (2016).

148. Zadeh, J.N. et al. NUPACK: Analysis and design of nucleic acid systems. *J Comput Chem* **32**, 170-3 (2011).

149. Oliveira, E. et al. Green and Red Fluorescent Dyes for Translational Applications in Imaging and Sensing Analytes: A Dual-Color Flag. *ChemistryOpen* **7**, 9-52 (2018).

150. Rosenthal, M. Clickmers and Aptamers as versatile tools for drug testing and fluorescence microscopy techniques. *Dissertation* (2020).

151. Toutchkine, A., Nalbant, P. & Hahn, K.M. Facile synthesis of thiol-reactive Cy3 and Cy5 derivatives with enhanced water solubility. *Bioconjug Chem* **13**, 387-91 (2002).

152. Wilson, C., Nix, J. & Szostak, J. Functional requirements for specific ligand recognition by a biotin-binding RNA pseudoknot. *Biochemistry* **37**, 14410-9 (1998).

153. Illangasekare, M. & Yarus, M. Phenylalanine-binding RNAs and genetic code evolution. *Journal of Molecular Evolution* **54**, 298-311 (2002).

154. Meli, M., Vergne, J., Decout, J.L. & Maurel, M.C. Adenine-aptamer complexes - A bipartite RNA site that binds the adenine nucleic base. *Journal of Biological Chemistry* **277**, 2104-2111 (2002).

155. Park, J.O. et al. Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nature Chemical Biology* **12**, 482-+ (2016).

156. Yamagami, R., Bingaman, J.L., Frankel, E.A. & Bevilacqua, P.C. Cellular conditions of weakly chelated magnesium ions strongly promote RNA stability and catalysis. *Nat Commun* **9**, 2149 (2018).

157. Park, J.O. et al. Near-equilibrium glycolysis supports metabolic homeostasis and energy yield. *Nat Chem Biol* **15**, 1001-1008 (2019).

158. Ronowska, A. et al. The Regulatory Effects of Acetyl-CoA Distribution in the Healthy and Diseased Brain. *Front Cell Neurosci* **12**, 169 (2018).

159. Sazani, P.L., Larralde, R. & Szostak, J.W. A small aptamer with strong and specific recognition of the triphosphate of ATP. *Journal of the American Chemical Society* **126**, 8370-8371 (2004).

160. Nutiu, R. & Li, Y. Structure-switching signaling aptamers: transducing molecular recognition into fluorescence signaling. *Chemistry* **10**, 1868-76 (2004).

161. Long, Y. et al. Selection of Aptamers for Metabolite Sensing and Construction of Optical Nanosensors. *Methods Mol Biol* **1380**, 3-19 (2016).

162. Ortega, A.D. et al. A synthetic RNA-based biosensor for fructose-1,6-bisphosphate that reports glycolytic flux. *Pre-print* (2020).

163. Autour, A. et al. Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. *Nat Commun* **9**, 656 (2018).

164. Tolle, F., Wilke, J., Wengel, J. & Mayer, G. By-Product Formation in Repetitive PCR Amplification of DNA Libraries during SELEX. *Plos One* **9**(2014).

165. A., S.L., Bontoux, N. & Stone, A.S. Formation of dispersions using ‘‘flow focusing’’ in microchannels. *APPLIED PHYSICS LETTERS* **82**, 364-366 (2003).

166. Mazutis, L. et al. Droplet-based microfluidic systems for high-throughput single DNA molecule isothermal amplification and analysis. *Anal Chem* **81**, 4813-21 (2009).

167. Bajic, A. et al. Relevance of the ability of fructose 1,6-bis(phosphate) to sequester ferrous but not ferric ions. *Carbohydr Res* **346**, 416-20 (2011).

168. Schutze, T. et al. Probing the SELEX process with next-generation sequencing. *PLoS One* **6**, e29604 (2011).

169. Pfeiffer, F. et al. Systematic evaluation of error rates and causes in short samples in next-generation sequencing. *Sci Rep* **8**, 10950 (2018).

170. Bailey, T.L. et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**, W202-8 (2009).

171. Ponchon, L. & Dardel, F. Recombinant RNA technology: the tRNA scaffold. *Nat Methods* **4**, 571-6 (2007).

172. Park, J.O. et al. Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nat Chem Biol* **12**, 482-9 (2016).

173. Macpherson, J.A. et al. Functional cross-talk between allosteric effects of activating and inhibiting ligands underlies PKM2 regulation. *Elife* **8**(2019).

174. Mookerjee, S.A., Nicholls, D.G. & Brand, M.D. Determining Maximum Glycolytic Capacity Using Extracellular Flux Measurements. *Plos One* **11**(2016).

175. Depaoli, M.R. et al. Real-Time Imaging of Mitochondrial ATP Dynamics Reveals the Metabolic Setting of Single Cells. *Cell Rep* **25**, 501-512 e3 (2018).

176. Xu, Q.W. et al. Can Galactose Be Converted to Glucose in HepG2 Cells? Improving the in Vitro Mitochondria! Toxicity Assay for the Assessment of Drug Induced Liver Injury. *Chemical Research in Toxicology* **32**, 1528-1544 (2019).

177. Whitney, J. Testing for differences with the nonparametric Mann-Whitney U test. *J Wound Ostomy Continence Nurs* **24**, 12 (1997).

178. Bennett, B.D., Yuan, J., Kimball, E.H. & Rabinowitz, J.D. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat Protoc* **3**, 1299-311 (2008).

179. Plaideau, C. et al. Overexpression of AMP-metabolizing enzymes controls adenine nucleotide levels and AMPK activation in HEK293T cells. *FASEB J* **26**, 2685-94 (2012).

180. Famulok, M. & Mayer, G. Aptamers and SELEX in Chemistry & Biology. *Chem Biol* **21**, 1055-8 (2014).

181. Yu, H. et al. In vitro isolation of small-molecule-binding aptamers with intrinsic dye-displacement functionality. *Nucleic Acids Res* (2018).

182. Rajendran, M. & Ellington, A.D. In vitro selection of molecular beacons. *Nucleic Acids Research* **31**, 5700-5713 (2003).

183. Szostak, J.W. Functional information: Molecular messages. *Nature* **423**, 689 (2003).

184. Raker, V.K., Becker, C. & Steinbrink, K. The cAMP Pathway as Therapeutic Target in Autoimmune and Inflammatory Diseases. *Front Immunol* **7**, 123 (2016).

185. Jenison, R.D., Gill, S.C., Pardi, A. & Polisky, B. High-resolution molecular discrimination by RNA. *Science* **263**, 1425-9 (1994).

186. Yuan, J., Xu, W., Jiang, S., Yu, H. & Poon, H.F.T. The Scattered Twelve Tribes of HEK293. *Biomed Pharmacol J* **11**(2018).

187. Johnson, K. et al. Mitochondrial oxidative phosphorylation is a downstream regulator of nitric oxide effects on chondrocyte matrix synthesis and mineralization. *Arthritis Rheum* **43**, 1560-70 (2000).

188. Zhong, G. et al. A reversible RNA on-switch that controls gene expression of AAV-delivered therapeutics in vivo. *Nat Biotechnol* **38**, 169-175 (2020).

189. Strobel, B. et al. A Small-Molecule-Responsive Riboswitch Enables Conditional Induction of Viral Vector-Mediated Gene Expression in Mice. *ACS Synth Biol* **9**, 1292-1305 (2020).

190. Tolle, F. & Mayer, G. Preparation of SELEX Samples for Next-Generation Sequencing. *Methods Mol Biol* **1380**, 77-84 (2016).

191. Beyer, H.M. et al. AQUA Cloning: A Versatile and Simple Enzyme-Free Cloning Approach. *PLoS One* **10**, e0137652 (2015).

192. Mathon, C., Bovard, D., Dutertre, Q. & Knorr, A. Impact of sample preparation upon intracellular metabolite measurements in 3D cell culture systems. *Metabolomics* **15**(2019).

9 ACKNOWLEDGMENTS

The first heartfelt thank you goes to Prof. Dr. Günter Mayer for allowing me to delve into the obscurities of SELEX, and sustaining my journey with a guiding hand and, at times, well-sought criticism. I owe much of where I am to that.

I want to express my utmost appreciation to Prof. Dr. Michael Ryckelynck for helping me build the foundation of this dissertation, and the digital biology of RNA team – Farah Bouhedda and Alexis Autour in particular – for the laughter and hard-work spent together.

I would like to thank my lab-rotation students, Shannon Fitch and Alexandra Shiian, together with the iGEM team members I supervised, for keeping up with me at different times of my doctoral studies.

My thank you also goes to Prof. Christoph Mayer and Dr. Marina Borisova for the fruitful collaboration, which finalized the last aspects of my dissertation.

Lastly, I would like to thank Prof. Dr. Irmgard Förster and Prof. Dr. Dieter O. Fürst for agreeing to review the present work and be part of the examination committee.

Throughout my doctorate studies, I was blessed to meet unique, single-minded individuals who truly amazed me with their passion, drive, and curiosity. Here some notable examples. To Fabian Tolle and Daniel Matzner, dearest friends and a constant source of inspiration: I am thankful for your invaluable support with the thesis correction, for having almost founded a company, and for the periodic Skype meetings, which I greatly cherish. To my darling Esther Volz, my sunshine on a rainy day, thanks for your help with the thesis, but especially for the absolute blast every time we are together! To Stefan Breuers, Cristian Renzl, Sebastian Pilsl, and Tjaša Legen for a great time spent together inside and outside the lab, thesis feedback, advice, and meaningful discussions, I salute you. To Moujab Choukeife and Georg M. Pietruschka, brothers from a different mother. We are, combined, the worst, amazingly dysfunctional, albeit perfect trio. I could not have wished for better companions. Thank you for the countless laughs, cries, hugs, yells, and messed up memories we shared together. It was not optimal, but it is our Ph.D.

A shout-out to Malte, Franziska, Silvana, Julia, and the rest of AKMayer, AKFamulok, and AKKath-Schorr, particularly to the PhD rookies Milena, Tejal and Rezzan: I wish you all the best for the future.

Vorrei ringraziare mia madre, mio padre ed i miei fratelli per l'amore giornaliero, ma soprattutto per la vostra assoluta certezza nelle mie capacità, che mi ha permesso di credere in me stesso anche nei momenti più duri. Ringrazio il mio padre spirituale Michele, le mie zie Enza ed Antonella, e mia cugina Giusy per avermi sostenuto nella preghiera, con quel aiuto che non si vede ma si sente. Grazie anche alla mia sorella, amica, e segretaria Cinzia, per la sua forza di spirito ed il suo aiuto concreto nel momento del bisogno. Inoltre, il mio enorme apprezzamento a Rosi e Federico; le mie batterie si ricaricano ogni volta che ci incontriamo.

Lastly, to Lisa and, of course, Kio. My emergency team. Learning to share my life with you felt as natural as breathing, and became just as vital. I love you with all my heart.